DOCUMENT RESUME

ED 185 008 **SP 015 722**

AUTHOR	Steffensen, James P.: And Others
TITLE	Teacher Corps Evaluation.
INSTITUTION	Nebraska Univ., Omaha. Center for Urban Education.
SPONS AGENCY	Office of Education (DHEW), Washington, D.C. Teacher Corps.
PUB DATE	78
CONTRACT	300-77-0156
NOTE	168p.
EDRS PRICE	MF01/PC07 Plus Postage.
DESCRIPTORS	Accountability: Educational Finance: *Evaluation Criteria: *Formative Evaluation: Government Role: Program Development: *Program Evaluation: Teacher Education
IDENTIFIERS	Teacher Corps

ABSTRACT

The papers selected to provide this resource deal specifically with evaluations of the Teacher Corps as opposed to evaluation of issues and problems generally. The first paper does address concisely the current dialogue in the educational evaluation community, but its role in this publcation is secondary to discussion of evaluation activities in the Teacher Corps today. The second paper offers a critique of ten years of evaluation of the Teacher Corps and also address the alternatives available to federal programs. It also discusses choices made by the Office of Education and Teacher Corps among those alternatives. The final paper is a statement of the work accomplished by the Evaluation of Teacher Corps Program 78 and Program 79. (Author/JD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

TEACHER CORPS EVALUATION

U S DEPARTMENT OF HEALTH
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

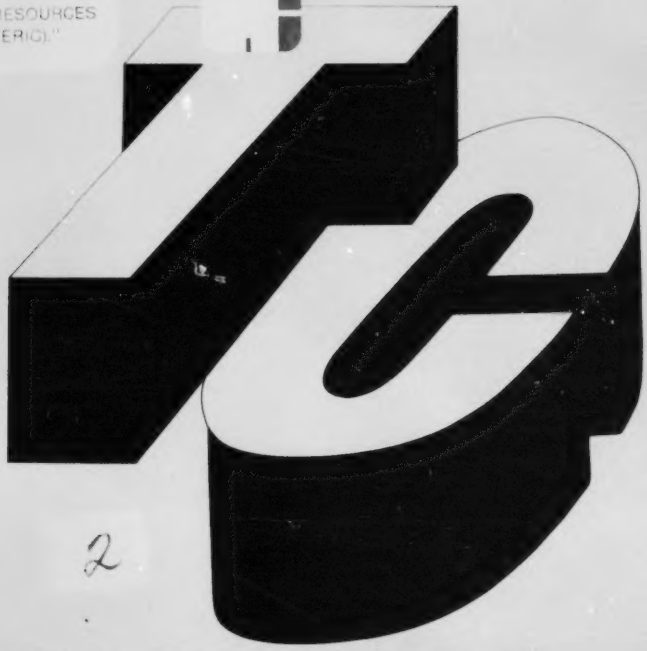
PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Theresa L. Kilgore

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ED185008

JAMES P. STEFFENSEN
G. THOMAS FOX
ROBERT BUSH
BRUCE JOYCE



2

10

TEACHER CORPS EVALUATION

James P. Steffensen
G. Thomas Fox
Robert Bush
Bruce Joyce

1978
TEACHER CORPS
THE UNIVERSITY OF NEBRASKA AT OMAHA

MAR 1 9 1980

CENTER FOR URBAN EDUCATION
The University of Nebraska at Omaha
3805 North Sixteenth Street
Omaha, Nebraska 68110

Project Director, Floyd T. Waterman
Editorial work by Arlene Low
Book design by Joan M. Krager

This material was produced pursuant to Contract Number 300–77–0156 between the Center for Urban Education at the University of Nebraska at Omaha and the Teacher Corps, United States Office of Education, Department of Health, Education and Welfare. Contractors undertaking projects under such government sponsorship are encouraged to freely express their professional judgment in the conduct of such projects. Points of view or opinion stated here, therefore, do not represent official positions or policy of the United States Office of Education.

Table of Contents

An Administrator's Perspective on Teacher Corps Evaluation 3

by James P. Steffensen

Limitations of a Standard Perspective on Program Evaluation: The Example of Ten Years of Teacher Corps Evaluations 11

by G. Thomas Fox, Jr.

Teacher Corps Evaluation Design: Excerpts from Task Forces Report and from Request for Proposals 87

Perspectives on 11th and 12th Cycle Teacher Corps Projects 119

by Robert N. Bush and Bruce R. Joyce

List of Figures and Tables

Figure

1 Comparison between Two Evaluation Perspectives 60
2 Comparision of Design Procedures of Two Evaluation Perspectives 66
3 The Generic Purposes of Teacher Corps 92
4 Preliminary Schedule 112
5 Illustration of New Way of Looking 165

Table

1 Comparison of Eleventh and Twelfth Cycle Projects in Selected Variables about the Nature of Evaluation 134
2 Focus on Evaluation Activities by Eleventh and Twelfth Cycle Projects 136
3 Classification of Evaluation Components of Eighty-Four Projects by Focus and Complexity 147
4 Classification of Evaluation Components of Eighty-Four Projects by Extent of Development and Implementation According to Focus and Complexity of Degree of Development 149
5 Data Collection Methods Dimensions by Phase of Project 152

An Administrator's Perspective on Teacher Corps Evaluation

JAMES P. STEFFENSEN

Introduction

The Teacher Corps Program is authorized under Title V of the Higher Education Act of 1965 as amended in 1976 (P.L. 94–482). Title V, Part B, Subpart I deals specifically with Teacher Corps. The specific purpose of the Teacher Corps Program is " . . . to strengthen the educational opportunities available to children in schools having concentrations of low-income families and to encourage colleges and universities to broaden their programs of teacher preparation and to encourage institutions of higher education and local educational agencies to improve programs of training and retraining for educational personnel by:

- Attracting and training qualified teachers who will be made available to local educational agencies for teaching in such areas;
- Attracting and training inexperienced teacher-interns who will be made available for teaching and inservice training;
- Attracting volunteers to serve as part-time tutors or full-time instructional assistants in programs carried out by local educational agencies and institutions of higher education serving such areas;
- Attracting and training educational personnel to provide relevant remedial, basic, and secondary educational training, including literacy and communication skills for juvenile delinquents, youth offenders, and adult criminal offenders;
- Supporting demonstration projects for retraining experienced teachers and teacher aides serving in local educational agencies. (Title V, Part A).

From 1966 through 1974 Teacher Corps has operated primarily with all but the demonstration focus. Grants are awarded to institutions of higher education and local educational agencies for the purpose of cooperatively developing and implementing projects to meet the particular requirements of the legislation as interpreted in the grant proposals. As of 1977 twelve funding cycles of the program had been initiated. Since 1972, appropriations for the program have been level at ap-

proximately $37.5 million. Grants to individual projects averaged $535,000 and covered a two-year period.

The most recent Teacher Corps Rules and Regulations published in the *Federal Register* on February 23, 1978, include four outcomes and several key program features. The outcomes stated in these Rules and Regulations are:

- Improved school climate which fosters the learning of children from low-income families.
- An improved educational personnel development system for persons who serve or who are preparing to serve in schools for children of low-income families.
- The continuation of educational improvements (including products, processes and practices) made as a result of the project, after Federal funding ends.
- The adoption or adaptation of those educational improvements by other educational agencies and institutions.

The key program features stated in these Rules and Regulations include:

- Education that is Multi-Cultural
- Diagnostic/Prescriptive Teaching
- Integrated Pre- and In-Service Training Designs
- Community-Based Education
- An Elected Community Council
- A Representative Policy Board
- A collaborative mode of operation involving the associated institution, communities and other vested-interested groups.

The Program has four major activity phases in time:

- Year I—Developmental Activities
- Years II and III—Operational Activities (pre- and inservice training
- Year IV—Institutionalization Activities
- Year V—Adaptation of Program Features by other schools, universities and other educational agencies.

All Teacher Corps projects are funded through competitive grants to *achieve the legislated purpose and the four outcomes using the means specified* in the Rules and Regulations and the strategies determined by the project. In general terms, the pol-

icy questions for this evaluation of Teacher Corps focus upon the relationships noted in the field between:

- the implementation of the key features as stated in the Rules and Regulations;
- the achievement of the four outcomes stated in the Rules and Regulations; and
- the realization of the purpose of Teacher Corps as authorized in legislative law under Title V.

Overview

The previous introductory section is an edited excerpt from the request for proposals which I have included so the reader might have an historical perspective on Teacher Corps itself before I turn to an overview of this volume on evaluation designs.

In late summer of 1978 the Office of Education will select a contractor who will conduct a study of Teacher Corps. This study will be the third major evaluation of Teacher Corps since its beginning some twelve cycles past. The other two—Corwin's analysis of Teacher Corps as a reform strategy and Marsh's study of the impact of Teacher Corps training on intern behavior—have provided many lessons for us. The study beginning with Program 78 has the potential, we feel, for equal if not greater impact upon Teacher Corps, the teaching profession and the education community at large.

The accessibility to and review of all appropriate data by all parties to an evaluation are prerequisites to the maintenance of a good information flow among the projects, the contractor, and the Office of Education. The papers which comprise this document have been selected to provide interested participants and observers with such information.

In June, 1978 the Office of Education issued a request for proposals for an evaluation of Teacher Corps Program 78 and Program 79. The statement of work was developed jointly by the Office of the Assistant Commissioner for Planning, Budgeting, and Evaluation and the Office of the Director, Teacher Corps.

The request calls for a major study of at least five years of

Teacher Corps, through 1983. It is to be a policy-oriented study for Teacher Corps/Washington, a study responsive to the local projects, a descriptive study of the actual practices taken by project and program participants and an illuminative study about the problems and difficulties, the degrees of success and failure and impact of implementing the Teacher Corps rules and regulations in field practice. In addition, the study is anticipated to be a learning experience for evaluating similar dynamic federal programs that have proven in the past to be difficult to capture and evaluate.

Teacher Corps provides a particularly fine case to which to apply such a policy-oriented evaluation. It is a thirteen year old program that has undergone and is undergoing major changes in direction. This means there are major policy questions that Teacher Corps management wishes to have answered with additional information. Teacher Corps has been evaluated often enough in the past to suggest that it is a difficult program to describe and evaluate, but it has also shown flexibility for change and responsiveness to evaluation results. Furthermore, Teacher Corps is a challenging program for evaluation since it addresses a common problem, that of improving the learning climate of schools for children from low-income communities, through a generally recognized intervention strategy: the training and retraining of teachers by the university, with a change strategy that reflects current professional wisdom. This means that an analysis of the difficulties encountered and practices generated from a Teacher Corps evaluation study should prove to be illuminative to other federal and state programs addressing similar social change issues, implementing similar intervention strategies and considering similar ideals for effecting institutional change. The evaluation study now just beginning should provide a grand opportunity to conduct such analysis. Teacher Corps has prepared this document as a resource for those individuals who will be most concerned over the next five years.

These individuals are the teacher, the parents, the deans, directors, principals, the professors—all who make up the project team. It is, in this sense, a recognition of the right to share in the ownership of an evaluation by those who are "being evalu-

ated", a recognition which is perhaps the single most distinctive and challenging feature of the current plans.

The papers selected to provide this resource deal specifically with evaluations of Teacher Corps as opposed to evaluation of issues and problems generally. The Bush/Joyce paper does address concisely the current dialogue in the educational evaluation community, but only secondary to the status of evaluation activities in Teacher Corps today. The Fox critique of ten years of evaluation of Teacher Corps also addresses the alternatives available to Federal programs and offers a reasoned criticism of the choices which the Office of Education and Teacher Corps have made to date among those alternatives.

A few pages of the Task Force Report on Evaluation are also enclosed. The close relationship of the suggestions made "from the field" in that report and the Evaluation RFP is, I feel, most heartening.

The final paper is the statement of work for RFP 78-79, the Evaluation of Teacher Corps Program 78 and Program 79. Each member of the project—directors, teachers, deans, parents, principals, interns—should become intimately aware of the proposals. The results of the evaluation will, we feel, have many implications for Program 78 and those to follow. More importantly, this evaluation has the potential to effect significant and necessary alterations in the role of the Federal government in the improvement of teacher education.

Teacher Corps Washington is appreciative of the contribution to its wisdom about program evaluation which Bush, Fox and Joyce have provided in the following papers. The work statement was essentially the responsibility of Eugene Tucker of OPBE and Linda Mount of Teacher Corps. Joyce Levine, of the South Dakota State Education Agency, and Reuben Olivares of the University of Texas provided very thoughtful insights into its early design. Patricia Kay of Baruch College, B. O. Smith of South Florida, and Robert Bush of Stanford each chaired one of the Task Force groups which helped shape our planning activities for Program 78 and whose joint statement so nicely stressed the necessity for strong interrelationships among the demonstration, documentation, evaluation, and technical assistance policies of Teacher Corps.

Limitation of a Standard Perspective on Program Evaluation

The Example of Ten Years Of Teacher Corps Evaluations

G. THOMAS FOX, JR.

(while you and i have lips and voices which
are for singing and to kiss with
who cares if some one-eyed son of a bitch
invents an instrument to measure spring with?

e e cummings[1]

Educational evaluators may understand this verse from cummings in two ways. First, it may reflect the feeling of many educators that program evaluation is unfair, unnecessary and a waste of time. Second, we may understand this verse as an indictment of certain practices that try to measure what cannot be measured or are too limited in their perspective for their own purposes.

The purpose of this paper is to address the second of these concerns, with the focus on federal program evaluations. Since there is seldom the option whether or not to evaluate, the challenge to the profession is how best to evaluate an intentional program that, like Spring, consists of change, life, growth, expectations and broken dreams. More specifically, the following question is asked: Are observed inadequacies of federal program evaluation due to failure to follow standard evaluation procedures, or are current standard evaluation procedures themselves inadequate for many evaluation purposes?

Although the question is not a new one, it has usually been discussed on either an ideological/theoretical level[2] or as a response to a specific large scale study such as the Coleman or Jencks report.[3] The discussion here will use as an example evaluation studies that have been conducted over the past ten years, on one federally funded program, Teacher Corps. Teacher Corps is a program initiated under the Higher Education Act of 1965 and is currently being funded at 37.5 million dollars a year.[4]

For purposes of this discussion, Teacher Corps should make a useful example for educational evaluators from a variety of

persuasions and interests. There are three main reasons why Teacher Corps should be of considerable interest.

First, the complexity of Teacher Corps has made it a special challenge to many different evaluators. During the ten years of its existence. Teacher Corps has undergone nine evaluation studies, from various perspectives. These perspectives have included studies funded through the Office of Management and Budgeting,[5] studies funded privately through the National Education Association and the Ford Foundation[6] and studies that have been initiated and sponsored by the agency itself.[7] Since Teacher Corps has been a challenge to evaluate, a number of creative procedures were used by some highly professional agencies.[8] In other words, the particular studies that will be reviewed in this paper have originated from a variety of sources and represent some interesting attempts at federal program evaluation.

The second reason that makes Teacher Corps an interesting case for evaluators is that it has continually shown a lively ability to be flexible in re-directing and changing its policy, focus and direction. Not only has Teacher Corps survived in an environment where most programs have not (at one time it was one of about ten programs associated with the Bureau of Educational Personal Development, now it is the lone survivor), but, judging from recent Presidential and Congressional advisory councils, it appears that Teacher Corps may grow in the near future.[9] Currently, in fact, it is undergoing a major shift in emphasis from preservice to inservice training and from the rhetoric of a service agency to that of demonstration projects.[10] Since Teacher Corps has been lively, flexible and dynamic in its policy changes, it is possible to compare policy changes to evaluation findings. This fact makes it possible to reflect upon the extent to which the evaluation studies have helped determine these changes.

In addition to the facts that Teacher Corps has been an interesting challenge to evaluators and that it has been relatively flexible in determining new policy and emphases, Teacher Corps may be interesting to a variety of educational evaluators because the field of teacher education is in need of more insight and understanding based upon empirical findings.[11] Teachers,

teacher educators and the general public have long agreed that there is much more to be learned about how teachers develop into better teachers. Thus, Teacher Corps is in the enviable position of being in a field where even a little new insight and illumination could be considered a dramatic increase in our understanding.

This paper will argue that the evaluation methodologies used, even in some of the more creative, insightful studies, were much too limiting, either for illuminating policy decisions or for furthering professional knowledge. The argument will be made by a brief review of the six evaluation studies of Teacher Corps, discussing their major findings, and relating these findings to Teacher Corps policy decisions and to the profession's understandings of the processes of teacher education. The studies will then be discussed using two different perspectives on evaluation, one representing a standard perspective,[12] one representing an alternative perspective.[13] After this discussion, a conclusion will be reached about the appropriateness of standard evaluation paradigms for federal program evaluations. Implications will then be drawn for future evaluation studies, with specific suggestions for changes in policy and procedure in evaluating federal programs.

Review of Studies

Six evaluation studies of Teacher Corps will be reviewed in this section. In chronological order, these six studies are:

- *An Assessment of Teacher Corps.* Report UR, 2 volumes. Bethesda, Md.: RMC Incorporated (Robert Crosby, director). August, 1970. Grant no. OEC-O-099018-4647. (Will be referred to as *RMC I.*)
- "The Teacher Corps: A Case Study in Evaluation," Chapter IV, in *The Role of Education in the Bureau of Educational Professional Development.* Boston: Center for Educational Policy Research, Harvard University. December, 1971. pp. 106-153. (Will be referred to as *Harvard*).
- "Teacher Corps," Volume I and Part G, Volume II of Report UR-172. *Full Scale Implementation of a Process Evaluation for Programs for NCIES.* Bethesda, Md.: RMC Incorporated (Robert Crosby, Director). November, 1972. (Will be referred to as *RMC II.*)

- Innovation and Change: A Study of Strategies in Selected Projects Supported in NCIES. 4 volumes. Cambridge, Mass.: (Will be referred to as *Abt.*)
- Corwin, Ronald G., *Reform and Organizational Survival: The Teacher Corps as an Instrument of Educational Change.* New York: Wiley-Interscience, 1973. (Will be referred to as *Corwin.*)
- *A Study of Teacher Training at Sixth-Cycle Teacher Corps Projects.* Volume I. Methodology and Findings. July, 1974. *A Study of the Effectiveness of Sixth-Cycle Teacher Corps Graduates.* Phase II. October, 1974. *Teacher Corps Graduates in Highly Innovative Roles in Education: A Follow-up Study.* Phase II. February, 1975. U.S. Office of Health, Education and Welfare, Office of Education, Office of Planning, Budgeting and Evaluation. Contract no. OEC-0-73-5174. (These three reports will be referred to as *Marsh.*)

The review of each of these studies will include:

- a brief introduction to each study including who sponsored the study, the purposes and major questions of the study
- a description of each study including the major variables identified, the data collection procedures used and the data analyses performed
- the major findings of the study
- the relation or impact of these findings to actual (or potential) Teacher Corps policy decisions
- the relation of these findings to professional understandings of processes of teacher education.

RMC I (1970)

Brief description of RMC I Study[14] This study was performed under contract to the Office of Education and was prepared specifically for Teacher Corps. The purpose of the study was to determine the extent to which Teacher Corps was achieving its stated goals in the field, and to attempt to identify program characteristics that contribute to success or failure. The major questions of this study were:

- What were the legislative goals and short term (process) objectives of Teacher Corps?
- To what extent was Teacher Corps achieving these goals and objectives?
- What were the structural characteristics of the program that contributed to success or failure in achieving these goals?

Study Design of RMC I. Since the major purpose of the RMC I study was to assess the degree to which Teacher Corps achieved its objectives, a discrepancy evaluation design was used. Several general legislative goals were identified and these were further broken down into forty-five process objectives that reflected Teacher Corps guidelines to funded projects. Then sample projects were identified and evaluated according to the extent to which the legislative goals and process objectives were reached.

Since a second purpose was to identify program characteristics that contribute to success or failure, six program "structural characteristics" were used to describe each program. Each of these program "structural characteristics" had two descriptive categories. This meant that there were six dichotomous terms that could be used to describe any particular program. The six "structural characteristics" or program variables were:

- program management—defined as either principally governed by the school district or principally governed by the university.
- team structure—defined as either "loose" where teacher trainees had no direct classroom responsibility, or "tight" where teacher trainees had direct classroom responsibility.
- program length—either 14 months or 24 months.
- program type—either graduate school training or undergraduate training.
- size of university—either small or large identified through enrollments in the School of Education.
- size of area served—either urban or rural and included the school district enrollment.

Although the ten sample programs were chosen in order to get as wide a representation of the six "structural characteristics" as possible, it was a limited sample that could represent only a fraction of the possible combinations of these six dichotomous variables.

Data collection involved a great amount of time and effort in identifying the forty-five process objectives through interviews and review sessions with Teacher Corps personnel. Assessments of project success in achieving each of these goals and process objectives were made through on-site visits, interviews with project participants, and questionnaires. For each project,

a description was drafted and a rating (of 3 possible levels) was given by two observers/interviewers. Differences in ratings and descriptions were resolved in confrontations between the two observers. Percentages of process objectives reached were computed for each legislative goal by project and for all the projects. Conclusions about the effectiveness of Teacher Corps were based on the percent of overall achievement for each of the seven legislative goals. Recommendations for improving Teacher Corps performance were made using these percentages of project effectiveness and the corresponding differences in the six structural characteristics.

Findings of RMC I. Teacher Corps had made clear progress in recruiting and selecting highly motivated student teachers. Furthermore placement of these student teachers in schools serving low-income areas did provide for small group instruction and for more noninstructional services to the community itself. There was little evidence, however, that the more ambitious legislative goals were being achieved. Individual Teacher Corps projects were having little impact on schools or universities. Local school practices were not being significantly changed and Teacher Corps training practices were being isolated and were not adopted by the cooperating universities. Although closer relationships were being developed between school districts, universities and communities during the life of the projects, there was little indication that this coordination would extend beyond the life of Teacher Corps funding. Community service was dependent upon individual efforts of the teacher trainees themselves and, thus, was disjointed and not significantly effective. Broader professional and public awareness of the accomplishments of local projects was not observed at all, not was there any evidence that others were being encouraged to adopt the successful practices.

Relation of RMC I findings to Teacher Corps policy. Recommendations from the RMC I study included a change in Teacher Corps priorities. At the time of the RMC I study, emphasis of Teacher Corps was to:

- train teachers
- provide services to the schools
- change university training

The RMC I study suggested this be realigned to:

- change university training
- change schools
- provide service to schools
- train teachers

There were no concrete suggestions on how to perform this change and, thus, there was little impact from these recommendations on Teacher Corps policy. (Later, as will be seen, the original Teacher Corps emphasis, as described in RMC I, was altered considerably.)

The other major recommendation from RMC I was to adopt the list of forty-five process goals that had been developed and to use these goals in an extensive internal management system for monitoring local project successes and failure. Again, since there were no concrete suggestions on how this would be actually accomplished (other than by questionnaires), it did not affect operational decisions of Teacher Corps. Five years later, however, a management system was incorporated into the Teacher Corps program.[15]

Relation of RMC I findings to professional understandings of teacher education. The arid nature of the six structural variables (e.g., graduate vs. undergraduate training) combined with the minimal ratings and description of the achievement of process goals make the illumination from this study extremely limited. Other than findings that Teacher Corps projects were having difficulties achieving the more ambitious legislative goals (regardless of individual energy and performance at the local level) there was little that would be considered to be an insight into teacher education.

Harvard (1971)

Brief description of Harvard study. This case study of Teacher Corps was part of a larger study on the Bureau of Educational Professional Development (BEPD). It was performed by the Center for Educational Policy Research under contract with the National Advisory Council on BEPD. The purpose of this study was to assess the role of evaluation in the operations of the program within BEPD. Teacher Corps was used as a sample

federal agency where the focus was on how Teacher Corps used evaluation in making, remaking, and changing its policy. There were three basic questions that the Harvard study tried to answer:

- Is evaluation of such programs possible? (And if so, how?)
- Is Teacher Corps the sort of program anyone in government wants to learn from?
- What does the experience of Teacher Corps suggest about the evaluation of discretionary, experimental and reform programs elsewhere in education?

In a sense, then, the Harvard study had a purpose similar to this report. It included describing Teacher Corps, stating what had been found about Teacher Corps and suggesting what may be useful for future Teacher Corps policy making and for future program evaluations. The difference between the Harvard study and this report, however, is that the Harvard study interpreted "program evaluation" as basically a singular procedure "resembling a research-development-field testing and evaluation cycle" and was seen as "*one way* of learning about the effects of educational reform and change." This report will be considering evaluation in a wider context and, in fact, will be suggesting an alternative perspective to the one assumed in the Harvard study.

Study design of Harvard study. This study was an historical, anecdotal report of how Teacher Corps was created by the legislature and how it operated as a decentralized federal agency. The purpose was to describe Teacher Corps official policy towards "evaluation" as well as the actual practice of Teacher Corps in using "evaluation." As stated before, the focus was on a certain perspective of evaluation and the question was to what extent evaluation as so defined was useful for Teacher Corps purposes. The irony is that the procedures used by the Harvard study to answer this question were not "evaluative," using their own definition.

The methods of data collection used in the Harvard study were difficult to ascertain but, judging from the spirit and content of the report, data were mainly collected through interviews, memos, anecdotal accounts of dinner meetings and lun-

cheon speeches, some informal reviews of legislative reports, Teacher Corps guidelines and an awareness of (at that time) ongoing Teacher Corps evaluation studies. Interviews with Teacher Corps/Washington staff were used.

The procedures for analyzing this information were also unclear but, again interpreting from the spirit of the report itself, were probably very informal, and idiosyncratic to the writers of the report. Little cross-checking or verification of the data seems to have taken place.

Reference to the data collection and analysis of this study must mention the liveliness of the writing style. The report is interesting, lucid and smoothly written, full of personal anecdotes, quotes from memos, legislative reports and proceedings and newspaper articles. The style invites serious mistakes of oversimplification and quite often seems to be written with a Freudian tongue in cheek ("Galbraith's ideas . . . may have been seminal but the Teacher Corps probabiy was born at a luncheon meeting in Washington, some months later"[16]). Nevertheless, the style also invites the reader's attention. The historical account of the legislative and administrative beginnings of Teacher Corps is very informative.

Findings of Harvard study. A general conclusion of this study was that Teacher Corps was essentially a political effort based upon long established assumptions about social change (e.g., reduce inequality through improved education). The political nature of Teacher Corps was indicated by continuous changes and revisions in its aims and methods. The study suggested that trainees were in contradictory roles of change agent, student and servant to the schools. It questioned the assumption that change in universities produces change in schools, and concluded that the key role for local projects was the school principal. The study emphasized the purse-string authority of Teacher Corps through proposal guidelines and grant making but concluded that local control of projects made it questionable that evaluative information could be generated or utilized. For example, it was critical of Teacher Corps' attempts to apply technical support systems such as "cumbersome" time-flow charts and cost-benefit analyses.

Generally, the Harvard study concluded that program evalu-

ation (as they had defined program evaluation) may not be possible in the Teacher Corps context. Although the rhetoric of Teacher Corps suggested it was an experimental program interested in program evaluation, in the view of this study, evaluation was of low priority and "one might almost say it is foreign to the character of the program itself."[17] This conclusion is found in the summary of the Harvard study. The summary begins by acknowledging that Teacher Corps (even by 1971) had existed longer than any other BEPD program. It then goes on to say:

> If any single point emerges from this study, it is the fact that the political and administrative climate in which Teacher Corps has been situated has not been particularly encouraging to evaluation. Before we conclude that Teacher Corps was incorrect in not supporting more evaluation efforts, or better ones, we must ask whether—given the organizational environment—it wasn't following a sensible course in paying r.o attention to evaluation.[18]

Relation of findings to Teacher Corps policy. As interesting as these findings are, they have had little to do with policy changes in Teacher Corps. Other than the rather obvious findings about the contradictory roles of interns, the importance of school principals and the question about the efficacy of the focus upon university change, there was little in this study that would relate to specific Teacher Corps policy. The conclusion that Teacher Corps was not interested in evaluation efforts is clearly false, judged by its current concern both with documentation and demonstration.

Relation of findings to professional understanding of teacher education. One may expect that the limited view of program evaluation and the rather flippant spirit of the report (e.g., that reform programs cannot be "evaluated") would restrict its usefulness to the profession. If, for example, the Office of Education held a view of evaluation less restrictive than that proposed by the Harvard study (i.e., if they held a notion of evaluation that depended less upon *a priori* objectives and predetermined treatments), then the major conclusions of the Harvard study would not be applicable.

Nevertheless, regardless of the flippant tone, inapplicabiltiy

to program policy, or faulty conclusions about Teacher Corps intentions, this report should be necessary reading for teacher educators. Teacher education is a political as well as an educational enterprise. The political negotiation that accompanied the original drafting of the Teacher Corps legislation, and that has continued to shape its existence and the definitions of its problems, is clearly a feature of *all* teacher education programs. As limited as this report is to specific program policy, the first section of the report throws much light on the political nature of teacher education, and provides some understanding of the extent to which shared notions of social change are consistent with designs for educational reform.

RMC II (1972)

Brief description of RMC II. The RMC II study was a two year study prepared for the National Center for the Improvement of Educational Systems (NCIES) and performed by the Resource Management Corporation. At the time, Teacher Corps was one of twelve NCIES programs. The major purpose of the RMC II study was to develop and then use a process evaluation system for all twelve programs. Although the aims of the RMC II study were to reach conclusions on funding/refunding decisions for all twelve NCIES programs, and to monitor all NCIES ongoing project and program activities, Teacher Corps did get special referral because it was different in many respects from the other twelve programs.

Study design of RMC II. There were two major variables used to compare the twelve NCIES programs: participant characteristics and project characteristics. Participant characteristics included the percentage of those recruited from outside the field of education, the percentage of participants expecting degrees, the percentage of those receiving state certification, and the percentage of those admitted to college on a provisional basis or for whom entrance requirements were waived. The project characteristics included the percentage of projects that changed state certification requirements, the percentage of projects budgeting for project continuation after withdrawal of federal funds and the percentage of projects serving schools eligible for Title I ESEA assistance, In addition to those vari-

ables identified for all twelve NCIES programs, there were also unique sets of participants and project characteristics for each of the twelve programs.

Data were collected by two surveys: one addressed the common characteristics of all NCIES programs and one was program specific. Both surveys were mailed to local program directors who were responsible for distributing the surveys. Survey data were cross-tablulated and descriptive data were reported in tables. A discrepancy analysis was performed using specific program goals and successful performance was determined by two types of summaries: one summary stated percent of measures successfully established by each project (this was used to rate each project), the other stated the percent of projects that met each measure successfully (this was used to rate each NCIES program).

Findings of RMC II. In comparison with other NCIES programs, Teacher Corps did very well. For example, most projects from other programs did not recruit from outside the field of education; Teacher Corps did. Most (75%) Teacher Corps interns received state certification; the average of the other NCIES programs was only 29%; most (88%) Teacher Corps projects served schools eligible for Title I ESEA assistance; the average of the other programs was only 48%.

The report of the project characteristics specific to Teacher Corps included recruitment, objectives, advisory councils and evaluation. Recruitment involved many different role groups and resulted in selection of trainees from a variety of racial/ethnic backgrounds. Objectives were focused upon changing university training procedures rather than improving or changing the school system. Although advisory councils existed, they seldom met. Almost all Teacher Corps projects said that evaluation was in progress, that measurable objectives had been set and that evaluation was being managed by the project staff (notice the contradiction between these findings and the predictions of the Harvard study). Formal evaluation activities (e.g., evaluation design, data collection and data analysis) were being conducted only by outside consultants and college staff. Informal evaluation activities (e.g., comments, opinions and perceptions) involved other project participants.

Other interesting survey results were that most (72%) participants felt that the local projects had caused them to do things differently. 30% felt they had new attitudes towards teaching, 33% felt that practical teaching experience was the strongest feature of their project and most felt that the weakest project feature was the lack of communication between participants.

It was reported that, for the discrepancy analysis, Teacher Corps had developed "the most extensive, detailed and comprehensive set of program conditions of any of the twelve NCIES programs."[19] RMC II also reported high minimum levels for satisfactory performance where over two-thirds of the projects equaled or exceeded 80% of the goals stated by the Teacher Corps program. The study concluded that these findings "indicate that projects are closely adhering to the majority of the program conditions, thus meeting the intent of the guidelines and the Congressional legislation."[20] Recommendations suggested for Teacher Corps included:

- increasing the flexibility of the interns' training program.
- placing specific emphasis on community participation and responsibilities of advisory council,
- taking action to ensure continuation of project developments,
- continuing the stress on practicum but placing more emphasis on the community component and,
- not letting interns substitute for classroom teachers or being responsible for regular classrooms.

Relation of findings to Teacher Corps policy. Since Teacher Corps was the only NCIES program to have survived, one could conclude that these findings were extremely important. On another level, however, the recommendations and findings have not been helpful for determining policy. The recommendations, for example, had no reference to context nor suggestions for procedure in performing the actions suggested. The figures for goal achievement gave no indication as to why some projects were achieving these goals and some projects were not. Thus, although the comparative results were healthy for Teacher Corps' existence and the discrepancy analysis suggested successful implementation of guidelines, the complete lack of evidence of the procedures being used to achieve

high performance left Teacher Corps with little to base future program policy.

Relation of findings to professional understanding of teacher education. The generality of the recommendations for NCIES programs and the rigid adherence to measuring only the perceived achievement of programmatic goals supplies no illuminative information to the profession at large.

Abt (1972)

Brief description of Abt study. This study was prepared for the Office of Planning, Budgeting and Evaluation, Office of Education by Abt Associates, Inc. and, like the RMC II study, was directed towards all NCIES programs (of which at that time Teacher Corps was one of eight). The focus of the Abt study was on the processes of institutional change, that is, how innovations are (or are not) implemented by institutions and incorporated into educational practices. The intent of the Abt study was to supply policy-relevant information useful in planning future "change-oriented" programs. The expressed purpose of the study was "to provide programs a framework for helping to determine the most favorable mix of project organization, innovation strategy and training content for reaching the desired goals."[21] The purpose was not to measure the impact of each project but to identify patterns within projects which facilitated or constrained change.

Some of the major questions of this broadly based study were:

- What are the constraints a project director faces and the strategies used to overcome these constraints?
- What are the influences that affect the effectiveness of programmatic aims directed through different institutions?
- What effects may be produced by variables that can be manipulated by the Office of Education, either through program guidelines or through grant award procedures?

Study design of Abt study. As may be expected from the wide-ranging purposes of this study, the study design was complex and involved. Initially, a conceptual scheme for describing project variation was developed. Projects were then examined

in order to verify and enrich that conceptual scheme. Using the scheme, strategies were identified that seemed most promising and productive. There were five major variations or dimensions (each representing a continuum) upon which a project was placed. These five variations included:

- the service orientation of the project (ranging from logical extensions of school structure to providing new personnel expected to shift school priorities),
- the training focus (from a university-based, nontraditional portal school environment to workshop programs organized outside the framework of the university setting,
- the practicum experience (from nontraditional roles for trainees to specific practice of particular skills),
- the centrality of the projects (from integrally built into the normal functioning of the university to separated from and involving little contribution from the university),
- the project's organization structure (from being diffuse in decision making to involving trainees and staff together in planning goals and strategies).

Data were collected through case studies. Thirteen sites were chosen as exemplary by their respective programs (two of these were Teacher Corps sites). Two interviewers (a "senior" and a "junior" interviewer) were trained. The training included a seminar on institutional change and conduct of a preliminary one-day structured visit which was then reported to colleagues. After the structured visit, a plan for a week-long visit was drawn and approved by the study director. Each plan included tentative hypotheses to be investigated, a full schedule of interviews for the first three days and a written confirmation of the interview schedule from the site.

All week-long visits were completed within a six-week period. Although the report designs of each visit were idiosyncratic to that visit, there were five content areas that had to be covered in all the case studies. These were:

- project descriptions, including verbal description of the institution of higher education (IHE), the local education agency (LEA), the community, and individuals,
- identification of strategies used to effect institutional change, including both "conscious" and "unconscious" strategies,

- the organizational dynamics for change including information as to who made decisions, an organizational chart of the project and the interviewers' judgment of how change could actually have occurred within the participating institutions,
- the impact of the project as perceived by project participants including the interviewers' own subjective interpretations of subtle areas of potential impact, and
- other unique features that might have been idiosyncratic to the project site and might also affect the replicability of the strategy. Several drafts, reviews and editings were made of each case study report with each site project director being involved in periodic reviews and editings of the reports.

Unfortunately, as extensive, open and detailed as these case studies were, they were reported but not actually used in the discussion of the results, nor in the recommendations. (This mistrust of the qualitative case study data may be attributed to the researchers' perception that case studies had never before been used in large scale evaluation studies.) Rating scales were used for cross-site analysis with the findings, conclusions and recommendations of the study being based upon the analyses of these ratings. Items for the rating scales were based upon literature on change theory and were written by the interviewers after their case study visits. Thirty-seven items were developed and each project was rated on the items using a Likert 7 point scale. (Items later were cross-indexed with the case study reports.) The thirty-seven variables were clustered by rater judgment, with the clusters being justified using a factor analytic procedure. Ten construct variables were then identified and construct means for each project were computed by transforming negative scores, adding all scores and dividing by the number of scales used in the construct variables. The analytic procedures included determination of means, standard deviations and pairwise correlations between all 37 scales and the ten construct variables, as well as partial correlations, multiple regression and causal analyses.

One interesting feature of study that was discussed but, unfortunately, not pursued in any detail, was a mapping sentence. The mapping sentence was designed to specify possible ranges of context and effects. It was a non-quantitative procedure using the judgment and perceptions of those constructing the

sentence. The general mapping sentence followed this format: projects of (_____), given contexts of (_____), cause institutional changes of (_____). Within each bracket were a number of possible descriptions. The "full mapping sentence" was more complicated than this general form because of further expansions of the descriptive categories within the brackets. The project characteristics were considered to be independent variables, capable of manipulation by the Office of Education or project management (e.g., amount of funding, duration of funding, pre-service or in-service). The context variables were certain characteristics of the participating institutions that could not be directly manipulated (e.g., attitudes towards change, support of a specific educational innovation). The change variables were the major concerns of both OE and project personnel. These included the permanence, magnitude and impact of the changes as well as where the changes occurred.

Findings of Abt study. The case studies on the two Teacher Corps projects were rich both in detail and in professional wisdom. For example, one of the case studies (a thirty-three page report) commended the project on its programmatic accomplishments but questioned its impact. The accomplishments were documented as new techniques for teaching reading and writing, new types of teacher recruits, modular instruction, team teaching, differentiated staffing and introduction of English as a second language. However, the report went on to call these changes "cosmetic," since there was no fundamental alteration in the problems associated with racism, educational bureaucracy or teacher-centered classroom instruction. The report further suggested that this problem may be endemic to the programmatic approach to educational change. A final section suggested that the project did nothing to change the tendency towards *de facto* segregation but, instead, seemed to "buy time" by defusing dangerous situations. The project did not deal with racial discrimination, community dissatisfaction with the status quo, maldistribution of resources or ineffective educational bureaucracies.

> Given the problem addressed, this Teacher Corps project has brought about the intended changes in the style and even the content of schooling in the classroom it serves. Yet, the consequences of

such innovations do not appear to be systemic. Although it may take four or five years to adequately measure the results of Teacher Corps community impact, new staffing patterns, curricula and concepts of instruction do not seem to deal with the primary problems faced by young children of low-income families.[22]

Compared to the tone and spirit of the case study report, the findings produced by the scales, construct variables and analytic procedures are relatively tame and uninteresting. The comparison between the two Teacher Corps projects and the other eleven projects showed many similarities in the 37 scaled variables except that:

- Teacher Corps personnel had more training,
- Teacher Corps projects had less institutionalization into the university and
- Teacher Corps projects had more risk and controversy than most other projects.

Probably the most important result of these comparisons, however, (and one not pursued in the study) were the large differences between the two Teacher Corps projects. The differences between the two Teacher Corps projects on all but one of the ten constructed variables were greater or at least as great as between either Teacher Corps project and any project from a different program.

Although all the results and findings of the Abt report that depended upon the construct variables and data analytic procedures must be highly suspect, there were three findings that were also corroborated in the case studies. These findings were:

- regardless of the value placed on change by the project staff or the presence of change training in the curriculum, it was the organizational structure of the project that was most closely related to implementing an innovation,
- the importance of "boundary personnel" within the cooperating organizations could not be overemphasized for change to occur and,
- the universities were extremely resistant to change.

Relation of findings to Teacher Corps policy. As rich as this report is in the description of two Teacher Corps projects and as con-

scientious in its attempt to specifically address policy decisions, there is little here that could be useful for Teacher Corps policy. The wisdom expressed in the case studies, for example, provides no clues about what could be done other than the ameliorative programmatic efforts already undertaken and accomplished.

Although the problems and paradoxes of programmatic funding were identified (e.g., technical innovation versus long-range institutional change), suggestions for alternative actions were scant. Simply put, this study was a good example that understanding the dilemmas inherent in pursuing certain goals was not the same thing as providing helpful suggestions.

The suggestions made, based upon the constructed variables, were both inappropriate and at best questionable as to their validity. On the other hand, the creativity with which the Abt study was pursued, especially the care with which the case studies were developed and, possibly, the sentence mapping technique for identifying manipulative programmatic variables may prove helpful to Teacher Corps in fulfilling its recent mandate for documentation, dissemination and demonstration.

Relation of findings to professional understanding of teacher education. First, the case studies of the Abt report should become standard reading to those interested in changing teacher education. They do illuminate the problems and provide concrete evidence of the dilemmas that confront a project funded for a finite period of time to produce specific innovations. Second, the data gathering and interpretive procedures mentioned above (the case study process and the sentence mapping technique), used with more confidence, may be helpful to those interested in documentation and evaluation of specific teacher education programs. Third, the review of literature on institutional change is extensive and should be informative to many organizationally illiterate teacher educators. Fourth, both the perception of and evidence for the importance of institutional and political arrangements when encouraging innovations in teacher education may be illuminating to teacher education. In short, although there are few specific recommendations that are useful to Teacher Corps policy, there is much in the Abt study that needs to be communicated to the profession at large.

Corwin (1973)

Brief description of Corwin. The Corwin report was a two-year study funded through National Education Assoication and the Ford Foundation. The study was headed by Ronald G. Corwin, a sociologist at Ohio State University and is the only Teacher Corps evaluation that has been published to date. The major purpose of this study was to gain insights into the possibilities for inducing institutional change. The study considered Teacher Corps as an interesting case of a federally funded program designed to produce institutional change. It used its findings to:

- generate some general propositions for subsequent research
- test other specific hypotheses from sociological theory and
- speculate about general issues of federal intervention, organizational change and educational improvement.

Unlike many of the other studies, the Corwin study was not necessarily designed to suggest specific changes in Teacher Corps policy.

The major questions of the Corwin study were:

- What is the role of the federal government in reforming this country's educational system?
- To what extent can complex organizations be deliberately changed?
- How can the professions be changed to improve service to low-income clientele?

The focus of the Corwin report is suggested by a quote from Sarason early in the study:

> In my experience, in practice, most explicit and implicit conceptions of change derive from the language and vocabulary of an individual psychology that is in no way adequate to changing social settings. The fact that one can be the most knowledgeable and imaginative psychoanalytic, learning or existentialist theoretician gives one no formal basis for conceptualizing the problem of change in social settings. The problem is simply not one to which these individual theories address themselves.[23]

Study design of Corwin. The assumption behind the Corwin study was simply that the capacities of organizations to change

must be enhanced before improvements in education could occur. Thus, the variables chosen were:

- variables related to organizational change and
- individual's attitudes towards change and initiative.

Thirteen variables were identified to indicate organizational characteristics or individual attitudes. Several of these characteristics described individual attitudes (e.g., commitment to teaching, initiative and compliance in education, role conceptions, alienation, liberation, satisfaction, preferred teaching styles). One described the extent of innovation by estimating the extent of change at each project due to Teacher Corps involvement. The final two variables described the organizational setting through a college quality index and a regional modernization index.

Data collection procedures included extensive use of questionnaires, a one-week on-site visit to ten projects by a team of interviewers (in 1968-69), and return visits to five projects a year later (spring, 1970). Selected for return visits were the more highly rated projects. Each site visit included interviews of personnel at three cooperating schools as well as one comparison school. Observation at the sites included field notes and field diaries as well as extensive interviews.

For most of the thirteen variables, multiple measurement scales and indices were used. These included a variety of Likert scale items and multiple choice items in addition to two expecially creative indices for measuring:

- initiative and compliance in education and
- professional and employee role conception.

Initiative and compliance in education were assessed using nine hypothetical incidents based on conflict between teachers and administration. Respondents (trainees) were asked to anticipate their behavior and estimate the seriousness of the incident. The two scales:

- what would be done and
- probable seriousness of the act in terms of official sanctions.

were combined to form a typology of initiative to compliance. Professional and employee role conception included four items to measure professional perspective. These four items were:

- regard for students,
- orientation to profession and colleagues,
- belief that competence is based on knowledge and
- belief that teachers should have decision making authority.

The employee scale included loyalty to administration and emphasis on rules and procedures.

The major dependent variable, inducements in organizational innovation, were measured according to assessments of Teacher Corps involvement in effecting changes in:

- school-college cooperation,
- classroom teaching techniques,
- school involvement in community and
- nature of college training.

Items were assessed using open-ended interviews and content analysis of planned and unplanned changes. A five point weighting scale was used on five dimensions (methods, materials used, clientele, personnel involved, and creativity of activity) for each planned change. Unplanned changes were judged by an additional dimension of improved attitudes. The ratings were based on the assumption that complex changes requiring new or altered role relationships are more important than simple additive changes affecting a few individuals. Each of the forty-two (42) schools had a scale score computed as a product of the number of planned and unplanned changes times the weights assigned to each applicable dimension.

The procedure also included a review of the literature on effectiveness of teacher training, teaching roles and effective teacher characteristics and arrived at an "ideal type" of teacher for poverty children. Comparisons were later made between this ideal type of teacher and the descriptive data on the trainees and practicing teachers.

A variety of data analyses were used to consolidate, compare and interpret the collected data. These included ANOVA tests

to compare variability within as well as between Teacher Corps projects, (e.g., on teaching effectiveness), correlations used to determine associations among dependent variables (to reduce number of independent variables), and regression analyses to explain the amount of variance due to newly constructed variables. Two case studies were also reported in order to show the different institutional contexts in which these projects were placed as well as to add more continuity to the discrete observations described above.

Findings of Corwin. The findings of this study reflected the broad spectrum of aims and the variety of information collected. Generally, the findings of the Corwin study gave a picture of the Teacher Corps trainees (interns), discussed the efficacy of the change strategies being used by local projects (with special focus upon the strategy of using the trainee as a change agent), described program outcomes in terms of university or school change, and discussed the theoretical implications for organizational change, with special focus on possible techniques for changing educational organizations through federal intervention.

The Corwin study, like the previous studies, found that Teacher Corps did attract different types of persons into teaching (for example, those with more initiative, less loyal to administration, more liberal). The trainees defined their roles and objectives differently from other teachers (stressing motivation, creativity and inquiry) and were more confident that children from low-income areas could learn. Although these trainees were not necessarily committed to a lifetime of teaching (in fact, Corwin later reports that they seemed both less confident and less appreciative of teaching classes on a day-to-day basis), they did approximate the proposed ideal type teacher more closely than did the experienced teachers. Furthermore, it was found that if trainees possessed one characteristic of the ideal type they tended to have all the others (i.e., the nine characteristics for the "ideal teacher" seemed to cluster for the trainees).

With these results, Corwin concludes that "probably the prime achievement of this program (Teacher Corps) was to provide a channel to attract this type of teacher."[24]

Evaluation of teaching competence was admittedly indirect and inferential, being constructed from personal ratings of cooperating teachers and supervisors. Four dimensions of good teaching were identified:

- innovativeness,
- understanding of pupils,
- rapport with pupils and,
- relationship with community.

Results of these ratings showed more variability of the ratings within projects and *between roles* within these projects, than between projects.

In reporting the findings on the teacher trainees, Corwin also reported some of his own primary beliefs on the nature of schooling. These included his belief that it is almost impossible to separate teachers' classroom behavior from their more general attitudes and values, and the related contention that a person's general style and personal values would make a greater difference in the outcomes of schooling than his/her techniques. The on-site visits made to classrooms convinced him that "most classrooms are the grim, joyless places described by Silberman."[25] However, Corwin also expressed doubts that the humanistic, ideal type of teacher would be appropriate or effective for low-income children (that is, for improving their national test achievement). He then goes on to suggest that the need might be for a greater blend of the academic and humanistic styles with a greater range of teacher models. Finally, in discussing the trainees, Corwin attributes most of the programmatic changes he observed to their optimism, their impatient criticism and their sustained pressure rather than to any other single program feature.

This brings us to the major findings of the study which related organizational constraints and strategies for change to the implementation of educational innovations. Since this is the major area of Corwin's expertise and, indeed, was the major focus of the study, it is these results which are the most interesting, trustworthy and informative. Corwin reports, for example, that contrary to a belief in the efficacy of professional collaboration, increasing teachers' active participation in the

project did *not* increase the likelihood of change. Indeed, participation by teachers was *negatively* correlated to:

- the introduction of new techniques,
- cooperation between college and schools and,
- the number of community projects undertaken.

Likewise, it was suggested that the professors in schools of education were unprepared to deal with the poor. They were too academic and remote from school problems.

In Corwin's view, Teacher Corps was applying two general strategies for inducing institutional change. One strategy was to rely upon technical authority and political power with little support of actual technical knowledge. The other strategy was to rely upon the power and influence of the inexperienced trainees as "change agents." Both of these institutional change strategies, Corwin concludes, were ineffective.

As reported in the Corwin study, the "trainee as change agent" strategy had its greatest impact upon the trainees. He reports, for example, that trainees became less disposed to exercising initiative, more disillusioned about school bureaucracy and administration, and less committed to teaching. The process of assimilation was described as being complex where some trainees (because of unwillingness to compromise or work within the system) removed themselves from a position of influence, other trainees (those who experienced least change) became "radicalized" out of the teaching profession during their two-year training, while a majority of all trainees became more convinced that basic changes in society were needed.

Teachers, on the other hand, showed a uniform decline in student orientation, an increase in loyalty to administrative rules and procedures, and an emphasis on pupil control. For example, technological innovations were negatively correlated to the teachers' attitude that society needed basic changes.

Innovative, substantive change (as measured by participants' opinions) was judged to be minimal at either the university or school level. At the universities, for example, the most that was anticipated were a few new courses, some modification of course content, employment of some new teaching techniques, (such as role playing, games and microteaching), an increase in

interest and attitude towards "disadvantaged students," and more dialogue between professors and students. Little improvement was expected in college-school cooperation (for example, no classes were to be placed on site) and no changes in the university or departmental organization were expected. Similarly, any changes anticipated in the school due to Teacher Corps involvement were short-range changes at best. Trainees assisted teachers by improving the student-teacher ratio, taking up some of the work load and providing individual attention to some students. There were few anticipated incorporations of new ideas, teaching methods, or materials at the school sites and, again, no expected changes in organizational structure and few changes in school-community relations. In short, Corwin concludes that the trainees only supplemented the normal activities of the schools and were used only as an added source of manpower by the teachers.

An assessment of the consequences of introducing liberal newcomers into the teaching profession produced the following interpretations.

First, the major dimensions of the environment contributing to innovation as the amount of cooperation between school and university. The combination of cooperating features that contributed to program goals was, however, far more complex than anticipated or imagined. For example, it was found that centralized decision making was more consistent with technological change than decentralized decision making. The control over funds, however, in these centralized organizations was also a leverage for co-opting the program. In addition, although centralized school organizations did not necessarily resist technological change, they did react defensively to critical, liberal, change-oriented newcomers. Finally, as noted previously, the belief in the efficacy of power equalization was further challenged by the negative effect that teacher participation had in inducing innovations.

Corwin cautiously summarized these results into some general propositions about how effectively to introduce innovations into school organizations. These propositions focused upon a combination of variables that could be expected to induce change more consistently and efficiently than the inde-

pendent change agent approach. The fundamental consideration, Corwin suggested, was the way local organizations were insulated from, or integrated into, a larger context. Within this discussion, Corwin concludes that future research should give more attention to the conditions that enable schools to maintain a functional autonomy within an increasingly interdependent, changing society.

When the dependent variable was the amount of school-college cooperation, regression analysis showed two general results. First, the critical ingredient was the quality of the "boundary personnel" or those who were assigned liaison roles between school and university. Second, cooperation between school and university declined with teacher participation in the proposal itself.

The final chapter (Chapter Eleven) in the Corwin study discusses these results within a theoretical perspective. The view is sociological, the topic is techniques for changing and improving educational organizations through federal interventions, the example is Teacher Corps. As a single piece of writing, this chapter is by far the most informative, creative and provocative forty pages of the thousands being reviewed in this essay. Chapter Eleven in the Corwin study should be required reading for all educators interested in implementing educational innovations.

Corwin interprets the main contribution of Teacher Corps as providing incentive and opportunity for more collaboration to occur among a variety of groups and organizations concerned with education. The critical innovation, from his viewpoint, is not curricular or instructional, but is the fact of mutual participation by schools, universities and communities. This view further suggests that the historical importance and significance of Teacher Corps is the deliberate establishment of a nationwide framework for collaboration among organizations to attack an educational problem (e.g., improving the education of children from low-income areas). Corwin uses Teacher Corps as an example for inferring an interpretation of the reform process itself.

Five categories of variables are suggested for interpreting the reform process. These categories are:

- the task structure of the organizations involved (e.g., goals, role structures and incentives),
- the status system and identities of members (e.g., background, authority, competence, values),
- economic resources available and how allocated,
- distribution of power and internal processes used to resolve conflict, and
- occupational setting.

Corwin suggests that of these five variables the distribution of power is the most critical, since it is often the distribution of power that determines the priorities of the other four categories. Innovation, he contends, will be received and implemented according to the results of the power struggles and related political processes of bargaining and compromise.

Two models of organization are then discussed. One is a rational model of organization, that is, that organizations are goal-directed entities where goals dominate the planning, the stating of alternative courses of action and the decisions. The other is a political model, that is, that organizational action is determined by external constraints, by previous commitments made in bargaining for resources and where official goals are only some of the possible sources of commitment and are themselves subject to compromise. Schools and universities, he suggests, can only be interpreted as political organizations. Negotiation, compromise, conflict and selective enforcement of conflicting norms are more closely associated with schools or universities than are the features describing rational organizations. The main point Corwin makes is that in a political organization, innovation must be *expected* to become ensnared in various internal and external power struggles that can only be understood in specific local circumstances.

With this in mind, the major variables that need be controlled by Teacher Corps, as an outside federal agency interested in inducing innovations, were reinterpreted. As interpreted by Corwin, these variables became:

- the sources of incentive and constraint (where goals become the bargaining units),
- an identification of the critical targets for change (for example, collaboration between schools and universities)

- a centralized guidance role for Teacher Corps where evaluations and systematic monitoring help determine funding through a flexible discretionary budget.

Lack of discretionary funds and the incapacity to gather information about local projects were suggested by Corwin to be major reasons that universities were so able to resist change.

Corwin concludes that Teacher Corps cannot be evaluated as if it were an agency that had a fixed, identifiable, achievable goal. Instead, faced with constraints beyond their control and with little concentrated power to implement a program as planned or even permit sufficient internal control, Teacher Crops should be assessed upon the more relevant and appropriate consensus-seeking functions. These consensus-seeking functions would be:

- sustaining the momentum for social reform,
- advancing alternative conceptions of innovation and
- facilitating debate and compromise.

Relation of findings to Teacher Corps policy. There are two ironies about the value of the Corwin report in helping determine Teacher Corps policy. The first irony is that, as provocative as the views and discussions of this study are, the few specific recommendations to Teacher Corps are relatively inane. One recommendation was that at least part of the intern training focus upon how to make changes in complex organizations. Another was selection of low-status universities and selection of a critical mass of trainees, and their placement in positions of control. Nevertheless, using the evidence of the Corwin study, one can easily go beyond the recommendations and offer the following suggestions. One, search for ways to include teachers in the training process (there is no way that teachers can be circumvented when inducing innovation, regardless of the difficulties noted in the Corwin study). Two, require specific institutional and role arrangements that can be expected to support a greater openness to and propensity for innovation. Three, emphasize the documentation, dissemination and demonstration of the methods used, to encourage both wider participation and greater consensus on educational issues. These are but a few of the specific recommendations that can be infer-

red from the Corwin study, which suggests its second irony. Although the Corwin study was the only one reviewed in this paper that had no intention of supplying Teacher Corps with specific recommendations for policy making, it has the greatest potential for helping advice.

Relation of findings to professional understanding of teacher education. Like the Abt report, the Corwin study should be required reading for teacher educators, both to acquaint them with a perspective that is necessary for understanding the process of inducing innovations in the schools and as a provocative example of the political and institutional constraints that influence efforts to implement educational reform. Unlike the Abt report, it is not the case studies that are so illuminating but the discussion suggesting a theoretical perspective on institutional change (Chapter Eleven).

It may be necessary to note that most of the findings of Chapter Eleven are not directly related to the statistical results of the study but are more related to the professional wisdom and interpretory powers of the author. As will be emphasized later in this essay, this inclusion of a personal and professional construct of meaning of the results can be interpreted as a praiseworthy asset rather than as a weakness of the study. Again, it may be appropriate to note that this is the one study that is published and is thus easily accessible to the profession. In addition, a volume of responses to the Corwin study by four teacher educators and a sociologist is also available.[26] If nothing else, this latter volume suggests the extent to which we teacher educators need be educated in the sociological perspective.

Marsh (1974)

Brief description of Marsh study. The Marsh study was performed under contract to the Office of Planning, Budgeting and Evaluation, Office of Education. Although it was begun under one corporation and finished under another, it was always under the control of its two principal investigators, David D. Marsh and Margaret F. Lyons. Possibly more than most other large-scale evaluation studies, it was also guided by the project monitor, Dr. Robert Hall of the Office of Planning, Budgeting and Evaluation. The aims of the Marsh study were

set very high. It was to be a study of intern recruitment, training and performance that could or would, once and for all, determine whether Teacher Corps training made any difference in the trainees and in the students they taught. The study was undertaken in the fall of 1973, and followed one cycle of trainees for the two-year training period and their first year of teaching. Up to now, it is the last effort to evaluate Teacher Corps performance by either the Office of Planning, Budgeting and Evaluation or any outside team of evaluators.[27]

There were two major goals of the Marsh study. These were to:

- Identify and analyze those combinations of intern background characteristics and Teacher Corps program characteristics that are related to desired teaching skills and attitudes of interns at the end of their training, and
- Assess the effectiveness of Teacher Corps graduates working with low-income/minority groups children.

The first goal was the focus of Phase I ending in the volume titled *A Study of Teacher Training at Sixth-Cycle Teacher Corps Projects*. In addition to the originally designed two-phase study, there was a follow-up component to Phase II that attempted:

- to re-analyze some of the data from Phase I and
- to describe educational activities that were involving Teacher Corps graduates other than public school teaching.

(*Teacher Corps Graduates in Highly Innovative Roles in Education: A Follow-Up Study*, is the volume describing other educational roles of Teacher Corps graduates.)

Before introducing the methodologies and findings of the Marsh study, it is wise to point out the considerable expense of the Marsh study, the lofty aims of its intentions, the great amounts of energy that went into its construction and also its place as benefactor from the previous attempts to evaluate Teacher Corps. Both its successes and failures must be understood in the context of a large-scale effort to judge the effectiveness of Teacher Corps intern training. One characteristic of this effort was the variety of data gathering procedures used and the multivariate data analyses applied.

Study design of Marsh study. Although the stated intentions were for Phase I to function in an hypotheses-generating mode and Phase II to function in further testing and refining of these hypotheses, the two phases actually had little in common. Phase I was designed to look for the most effective ways to train interns. Generally, this was done by searching for relationships between and among:

- trainee background characteristics,
- training program characteristics, and
- trainee classroom performance characteristics (called "exit characteristics").

Three of the goals of Phase II, on the other hand, were designed to assess the effectiveness of Teacher Corps graduates in their first year of training. Generally, this assessment was approached by contrasting the performance of Teacher Corps graduates (and their students) with that of control group (non-Teacher Corps) graduates. Although a fourth goal of Phase II was to relate background, past training, teacher performance and pupil performance, the only relationship between this goal and Phase I was to include the data describing the program characteristics for various canoninical and multivariate analyses with later performance data. There was no revision, clarification or further testing of the hypothetical statements made in Phase I from correlational analyses. Simply put, the Marsh study was really two studies (with the third volume reviewed here as an unrelated account of interns' experiences after graduation).

In addition, any underlying rationale or theoretical position of either phase had to be inferred. Phase I, for example, chose to define the teacher training process in terms of three major variables:

- intern background characteristics,
- program characteristics, and
- performance or "exit" characteristics.

Although the input, output, process approach to defining the problem may be obvious when given these three variables, it was never stated, nor were results from studies of similar per-

spectives cited. There was, for example, no attention given the differences often noted in such studies between "manipulative" variables (those that cannot be expected to be changed). In simple terms, both Phases of the Marsh study were designed to produce *tour de force* statements about the effectiveness and impact of Teacher Corps training upon its recipients.

This superhuman aim was reflected in some almost superhuman efforts at data gathering through observations, interviews and questionnaires distributed through repeated on-site visits.

In Phase I, data were collected from all 20 Teacher Corps projects preparing elementary school teachers. Performance or exit characteristics were collected on sixty percent of the interns in a stratified random sample (by age and sex). Site visits of from one to two weeks were conducted at all sites both in the fall and spring semester. Training for these visits required nearly three weeks. There were eleven questionnaires and about seven interview schedules used to gather data on the program characteristics above. Four instruments were used to "measure" performance or exit characteristics. These included a direct classroom observation guide, a guide for reconstructing trainee activities (both a questionnaire and interview schedule), supervisors' (team leaders) assessments of specific abilities and activities of each trainee, and a trainee self-rating form. The background data were collected through a thirty-item questionnaire.

The amount of time, effort and concern given the data gathering procedures for Phase I was, obviously, considerable as were the amounts of data collected.

Data analyses for Phase I consisted mainly of factor analyses to reduce the number of variables. New program "factors" and performance (or "exit") factors were constructed using orthogonal and oblique rotations. Canoninical correlations were performed on the background, program and performance variables to determine "important trends" between program variables, background variables and performance. Multiple linear regression analyses were used to determine relationships between background and program factors with specific performance or "exit" characteristics. Relationships were drawn for

the sample population of trainees as well as for three ethnic groups of trainees (Black, Chicano and White). Two-way analyses of variance and *t*-test comparisons were used to determine differences between graduate and undergraduate training projects on performance skills. Simple correlations were used to search for the extent to which Teacher Corps projects had had an impact on their host universities implementing a competency-based teacher education program.

For Phase II, a sample of Teacher Corps graduates (interns) was taken from the original pool of trainees and compared with a sample of recent graduates from non-Teacher Corps training programs. Both samples were teaching in public elementary schools (grades 2-6) with major responsibility for a specific group of children. Individuals from each sample were matched on grade level, pupil socioeconomic background, school district and years of teaching experience.

Performance variables upon which the two samples were to be compared were modified from the performance factors of Phase I. Twenty-three performance factors were reduced to seven teacher performance clusters. These seven performance clusters were:

- utilization of community resources,
- action as a change agent in the school,
- lesson planning and materials development,
- classroom performance: affective tone,
- classroom performance: teacher response strategies,
- individualization of instruction, and
- perceptions of problems related to schooling.

Data on teacher performance were gathered by a teacher questionnaire developed for the Marsh study and by use of a previously developed classroom observation guide.

As stated before, the aims of Phase II were high and included comparison between Teacher Corps and non-Teacher Corps graduates not only on classroom behavior but also on their pupils' performance. The variables upon which students were measured included:

- reading gain,
- self-concept gain,

- school attendance, and
- pattern of pupil classroom interaction.

Data on these pupil variables were collected using the Metropolitan Achievement Test for reading ability, the Piers-Harris Self Concept Scale, school attendance records and the previously mentioned classroom observation guide for pupil interaction.

In addition to teacher performance variables and pupil variables, Phase II also identified school-climate variables that were considered as possible influences on teacher performance and effectiveness. These variables included:

- the percent of low-income pupils,
- feelings of the teacher about other staff members, and
- the school setting.

This information was collected by the previously mentioned teacher questionnaire.

All data for Phase II of the Marsh study were collected in the fall and spring of the 1973-74 school year. The reading achievement and self-concept tests were given both in the early fall and late spring of the school year. Classroom observations were conducted and teacher questionnaires distributed in the spring. There were twenty five-minute observations taken of each classroom over a consecutive two day period.

As in Phase I, Phase II used a variety of data analytic procedures to "answer" specific questions. For example, one-way Analyses of Variance were used to compare the two samples of teachers (the Teacher Corps graduates and control group) on both the teacher performance measures and the pupil measures. A one way analysis of covariance was used to compare the two samples on pre- and post-scores from the reading and self-concept tests. Canoninical correlations were used to identify teacher performance variables with pupil variables. Discriminant analyses were used to identify those teacher performance variables which best predicted "above-average" or "below-average" classroom mean scores. Simple correlation and discriminant analyses were used to search for relationships between teacher sex, teacher ethnicity and Teacher Corps pro-

gram characteristics with both teacher performance variables and pupil performance variables. Simple and canoninical correlations were used to determine relationships between school climate variables, teacher performance variables and pupil performance variables. A one-way analysis of variance was used to compare Teacher Corps and control group samples on each school climate variable. A Chi-Square analysis was used to compare the Phase II sample of Teacher Corps graduates with all interns from the Phase I study on sex, ethnicity and project affiliation.

The follow-up study of Phase II was principally a descriptive account of Teacher Corps graduates' activities (and intentions) in the year following graduation. These descriptions were achieved in two ways. One was through a questionnaire sent to all Teacher Corps graduates (669) prepared as elementary school teachers. Approximately 50% (338) of these questionnaires were eventually returned. The questionnaire focused upon rudimentary job descriptions (e.g., extent to which schools served low-income pupils), trainees' reflections upon their Teacher Corps training experience (e.g., quality and usefulness of training to current position), trainees' attempts to locate jobs and their plans for the future. Results from the questionnaire were given in terms of the percent of those giving each type of response (e.g., 79% reported they worked in schools that served low-income children, 42% reported that their Teacher Corps training was "extremely useful").

In addition to the mailed questionnaire, there was an hour-long phone interview with approximately 20 graduates who had "innovative roles in education." These roles included:

- teachers of pupils with learning disabilities,
- teachers of English as a second language,
- reading, math and science specialists,
- special non-teaching jobs in education, and
- jobs in institutions of higher education.

Three to five graduates from each of these five job categories were selected on the basis of their job description in the questionnaire. The interviews were conducted in November and December, 1974. Although the expressed intention of the in-

terviews was that "each job description would be somewhat independent of the unique experience of any one graduate,"[28] the method used to come to these general job descriptions was not described. Descriptions of each job included:

- general statements about task responsibilities,
- how those in that job category related their Teacher Corps experience to their current responsibilities.

The number of graduates interviewed for each job category (three to five) and the purpose for these descriptions ("to identify patterns in graduates' professional growth"[29]) left these job descriptions with a dilemma, how to generalize from three to five cases when not reported in detail. That the resolution of this dilemma was not successful is clearly in evidence from the many times that phrases similar to "all teachers reported"[30] were stated when "all" referred to three, four or at the most five individuals. One time, for example, the statement was "Several teachers found their Teacher Corps coursework helpful, several did not"[31] when there were only four persons who were interviewed. This attempt to make the job descriptions general rather than context-specific lessened the potential value of reporting and analyzing the phoned interviews.

These three separate phases of the Marsh study, taken together, represented a serious, professional attempt to produce information useful to Teacher Corps and to the teacher training profession. Not only was there an extensive use of surveys, questionnaires, interviews, observations and on-site visits but also all these techniques (with the possible exception of the case studies) were approached with a high regard for instrument validity and observer reliability. Furthermore, the data analytic procedures applied were some of the more sophisticated techniques available for correlation data. Although this reviewer has elsewhere criticized the categories used for identifying many of the variables in the Marsh study, and the lack of a theoretical base,[32] any consideration of the Marsh study must include great respect for the amount of effort and diligence that went into gathering and analyzing data. Compared to the other studies reviewed in this paper, there is no doubt that the Marsh study was more ambitious in its intentions, more

energetic in its approach and more conscientious in fulfilling its evaluation responsibilities as established by most current definitions.[33]

Findings of Marsh. Considering the zeal with which data were collected and the technical expertise by which the data were tested and analyzed, the irony of the Marsh study is the lack of illuminative findings. Admittedly, most of the evidence supplied through the Marsh study was correlative, that is, "strength of relationships" between certain variables and constructs of variables. Nevertheless, realizing that correlation data are never designed to give causal relationships but are, rather, designed to suggest whether certain relationships do or do not occur within certain contexts, the results are still relatively tame and uninteresting. As empirical as the Marsh study clearly was intended to be, most of the findings of Phase I can be interpreted as definitional problems.

For example, one of the "most important correlations"[34] between program factors and performance (or exit) factors was the strong inverse relationships (−.60) between the use of teacher competencies in the training program and the extent to which trainees felt that they were self-directed. The use of teacher competencies in training meant:

- that certain skills were pre-determined as being necessary to successful completion of the training program,
- that these skills were determined before the training began,
- that all training activities were directed to achieving these skills,
- that mastery of these skills was assessed periodically by predetermined means and
- that as soon as a trainee passed the assessment, he or she could move on to the next skill.

Considering that this describes activities underlying the phase "the extent that competencies were used in training," the interesting result would have to be that the inverse relationship was "only" .60. Why would trainees feel self-directed in such a system? The answer to *that* question would have been illuminating. Partial answers might be:

- the trainees were duped into feeling self-directed,
- the use of competencies was not as efficient in practice as was defined and intended by the trainers *or*

- there were other activities in the interns' training that countered the determinism implied by the use of stated competencies.

Here empirical evidence could have been informative, rather than merely a continuation of a logical relationship.

A similar example was the "finding" that the follow-up interns' academic instruction (i.e., the extent to which university course work was tied into later classroom activities) was related to teacher participation in the project and to cooperative decision making between teachers and university professors. This finding, again, seems merely tautological where it would be far more illuminating to find out how follow-up was done *without* teacher participation or cooperative decision making (the relationships were .57 and .44 respectively). Empirical evidence would have been interesting only to the extent to which it did *not* corroborate the logical relationships.

Not all findings of Phase I of the Marsh study were of this variety. Some of the correlations tell what is already common knowledge. For example, positive attitudes held by university personnel towards a training program are inversely related to the extent to which teachers participate in the project (the old two-culture or practitioner vs. theoretician problem) and the size of the school district is inversely related to the trainee's opportunity to bring about change in the school.

Nevertheless, there are a few results that warrant attention because they are difficult to explain or to understand or because they seem to contradict present professional wisdom. For example, the trainee's perception of the extent to which they have changed schools and the importance of these changes is inversely related to the amount of self-initiated contact with parents and to children-initiated interaction within the classroom. This suggests that the trainee who is more autonomous and dogmatic in the classroom is the one who perceives her/himself as bringing about important change. Those who may be most innovative (e.g., who initiate more contact with parents and who allow children to initiate classroom interaction) give less importance to school change. The irony of this finding (i.e., those who make most changes see futility in these changes, those who make smaller changes see the changes as greater

than they are) and the potential conflict this may raise within and between individuals should not be forgotten or easily disregarded by teacher educators.

Perhaps the clearest challenge to our professional beliefs comes from the relationships (more specifically the *inverse* relationships) found between trainees' performance characteristics. Before nothing these relationships, it is appropriate to be reminded that these performance characteristics are not only the most clearly defined variables identified in the Marsh study, they also have been "measured" by some of the most reliable means used in this study. The characteristics were both recorded by trained observers and/or were self-reported as actions (rather than attitudes) by participants. There were twenty-nine performance characteristics (e.g., degree that instructional choices are given to pupils, intern initiates contact with parents, etc.). The relationships between many of these actions are surprising and beg for further interpretation. Some of the more surprising correlations are:

- trainee use of a broad range of resources for preparing lessons was inversely correlated to trainee's contact with parents and also inversely correlated to trainee's acceptance and use of student ideas, whereas,
- the same characteristic (use of broad range of resources) was positively correlated to the use of an informal authority structure in the classroom and the extent to which trainees provided different learning activities for different children, furthermore,
- the trainee use of an informal authority structure was inversely related to attentiveness to children and also inversely related to an ability to communicate with pupils. In other words, the more trainees accepted and used student ideas, the fewer resources were used for instruction.

The more informal the authority structure, the less trainees accepted student ideas, but the more they used a greater variety of resources. In short, in situations that teacher trainees usually identify as especially humanistic and capable of flexibility (informal rather than formal authority), they seemed to behave more inflexibly and interact less with the students. In the Marsh report, there were many such findings about trainee performance that seem to defy current assumptions or supposedly logical relationships between behaviors.

Another example of these provocative findings were the strong inverse relationships between the trainees' following lesson plans and their use of other skills that one would have expected to be related to lesson planning. Effective pupil diagnosis, the extent to which trainees provided different learning activities for different children, use of open-ended questions, attention to children's responses and praising children all were negatively correlated with following lesson plans. In other words, instead of lesson plans opening up the possibility for more appropriate responses to students, they seemed to cut off the trainee from the students.

There were many such interesting results of the correlative relationships between certain trainee performance activities. Although these results can be interpreted from different perspectives (and, in fact, beg to be interpreted), there was no attempt in either the Phase I or Phase II reports of the Marsh study to deal with some of the more surprising or seemingly contradictory results. The one thing that was done, however, was to suggest two different "prototypes" of trainees. Type A introduced cultural relevant materials, did not follow lesson plans, did not allow children to explore the room or select a work group. Type B focused on academic topics, allowed students to move about and select work groups, used lesson plans and used corrective feedback. The incongruity of the characteristics within each type was not addressed.

The remainder of the findings of Phase I tried to determine relationships between trainee background, program characteristics and intern performance characteristics. Generally this proved to be unsuccessful. Predicting trainee performance characteristics from trainee background and program characteristics, for example, suggested only that one could predict a trainee's perception of the causes of poverty and poor reading skills. (These in turn were related to low-income and minority setting of the program.) Likewise, relating program factors with performance factors also provided little information. Out of sixty-five potential program factors, nineteen were correlated with at least one performance variable.

Generally, then, other than point out that little could be predicted about trainee performance from program or

background characteristics, Phase I failed in its original intent to "identify and anlayze those combinations of intern background characteristics and Teacher Corps program characteristics that are related to desired teaching skills and attitudes of interns at the end of their training."[35]

Phase II, however, one which set its goals even higher (that is, to assess the effectiveness of Teacher Corps graduates in working with low income/minority group children) was actually more successful in terms of that goal. Some interesting differences (and similarities) were found between graduates trained by Teacher Corps and graduates trained elsewhere (the control group).

No differences were found in classroom performance characteristics. There were, however, three out-of-class performance characteristics that did differentiate Teacher Corps graduates from other graduates. Teacher Corps graduates:

- developed more ethnically relevant curricula and
- used more community resources in teaching and
- initiated more contact with parents.

There also was a difference in attitudes towards the causes of poverty and the causes of reading problems, with Teacher Corps graduates attributing these causes to society and environment, other graduates attributing the causes to the individual.

There were no differences noted in pupil reading achievement. There were significant differences, however, in pupil self-concept scores. The pupils of Teacher Corps graduates had higher scores of self-concept about their physical appearance and higher scores of happiness and satisfaction with themselves. In summary, no differences in classroom interaction and teaching styles were observed between Teacher Corps and non-Teacher Corps graduates; no differences in pupil achievement were observed, but there were differences in out-of-classroom activities (curricular design and use of community resources) and differences in pupil gain on certain self-concept subscales.

Teacher Corps graduates were also found to perceive that their views about curriculum and discipline were more diver-

gent from other teachers than control group graduates. Likewise, Teacher Corps graduates were less likely to receive ideas and curriculum materials from other staff members. Generally, it was found that for both Teacher Corps and non-Teacher Corps graduates, those who "individualized" instruction (i.e., allowed children to work in small groups and/or individually and praised the children a good deal) were also likely to be those who:

- did not obtain ideas or materials from other teachers and
- perceived their views as being not similar to other teachers.

Considering the difficulties inherent in managing individualized instruction, this combination of attitudes might be considered to be self-defeating for inexperienced teachers in self-contained classrooms.

The attempt to relate training program features with graduate performance and pupil performance was not as successful. The few relationships found between program features and graduate performance features were the positive correlations between:

- low team leader-trainee ratio and the use of community resources and
- strong team leadership during training and involvement in a change agent role in the schools.

The few relationships between program features and pupil gain scores were even less informative (e.g., higher gains in pupil self-concept were related to projects with courses taught by white instructors).

Finally, correlation analyses between the teacher performance characteristics and pupil gain scores suggested that large pupil gains in both reading and self-concept were most closely related with:

- teacher perception that reading development was primarily a function of pupil characteristics rather than teacher characteristics
- a low level of teacher/pupil interaction
- teachers acting to bring about change in the school,
- teachers initiating contact with parents and
- planning lessons with a team of teachers.

Findings in the follow-up portion of the Marsh study suggested that the majority (about 88%) had sought teaching jobs, had jobs in education (85%), mostly as teachers (77%) and were in schools serving low-income children (80%). Most planned to continue as classroom teachers in the following year (77%) and all who had jobs in education planned to remain in education. Most planned to stay in education after five years (82%) but, in general, saw themselves moving from classroom teaching (from 77% to 31%) to other educator roles (from 8% to 51%). Most rated highly the quality of their Teacher Corps training as well as the usefulness of that training to their current work.

Although the case studies of those in innovative educational positions were not formally summarized, the following were some generalizations made in the Marsh study. The more common themes stated about Teacher Corps training experiences were:

- the importance of the trainees' relationships with their training team leader
- the importance of seeing a variety of instructional techniques
- the value of gaining extensive experience in public schools, not necessarily in teaching but in understanding the operational aspects of schooling
- their sense of personal growth during training, especially in dealing with frustration and ambiguity as well as taking personal responsibility for initiating action,
- the sense of commitment to educational innovation developed within a project related to a national effort,
- the problems within their training project, (for example, little support and much disorganization actually forced them to take initiative and organize),
- the value of their community involvement,
- the view that only courses directly applicable to their special area of responsibility were useful, and generally,
- that their training experience was mostly protrayed as a "consciousness-raising experience, broadening their notion of education, schools, and teaching."[38]

A final theme of those in educational roles other than teaching was the energy toll of their work. Many were reported as feeling weary, isolated and overworked. Many felt a need to

share their accomplishments with others but found little opportunity to do so.

Relationship of Marsh study to Teacher Corps policy. There are two important points to make about the relationship of the Marsh study to Teacher Corps policy. First, is the rather unfortunate circumstance that the completion of the Marsh study, directed at assessing the effectiveness of trainee preservice training, coincided with a change of Teacher Corps from primarily a training program for certifying new teachers to a training program for experienced teachers. Teacher Corps, after the Marsh study was completed, was quite different from the Teacher Corps evaluated by the Marsh study. Thus, one would expect that the Marsh study would not be particularly valuable as an aid in determining future Teacher Corps policy except as it may relate to the continued preservice component of Teacher Corps projects. (Four instead of thirty-five trainees are trained for certification while an entire school staff is the focus of the re-training component.)

The second point, however, is that the results of Phase II of the Marsh study (that there were *some* differences that favored both the performance of Teacher Corps graduates *and* their impact on students) have been effective in supporting Teacher Corps at a time when decisions were being made about its survival. The directness of its findings and the simplicity of the Phase II design made the findings of the Marsh study valuable to Teacher Corps' current status in the Office of Education. To those evaluating Teacher Corps, the Marsh findings suggested that:

- Teacher Corps was doing what it was supposed to (recruit and train minority members to teach in low-income schools),
- that its graduates were entering the educational field, primarily as teachers and
- that this training was having some impact upon the pupils in these schools.

These results suggested to policy makers that, even though major changes in Teacher Corps emphases were being made, it may be able to carry off these responsibilities as well.

Nevertheless, when considering these two points about the

Marsh study, one must still conclude that, as the study now stands, there is relatively little of value to Teacher Corps policymakers. Although the quantity of data collected in Phase I was great, there is little interpretation of these data that can be used to identify possible agency alternatives in training nor even to judge whether certain training procedures are or are not effective. Potentially, the most useful data (and certainly the most reliable) are those describing trainee performance characteristics. Unfortunately, because of the grand design to determine relationships between program features and impact, these data, especially their ambiguity and incongruity, became relatively unimportant to the study. The contradictory results of the Marsh study, pursued further, could suggest some interesting and illuminating consequences of certain training procedures used within the Teacher Corps context. Incongruities such as:

- contacting parents but
- not developing culturally relevant curricular materials, or
- following lesson plans but
- not providing different learning activities for different children

have not been anticipated nor documented before.

To resolve these incongruities would be to address specific training procedures. For example, it could be suggested, since doing each activity well takes so much time, each training program should allow different trainees to focus on different features, or that there are needed other training strategies that better address and combine these observed incongruities. It must be noted, however, that the Marsh study makes no mention of such possibilities. In fact, one could anticipate that, because of the lack of interpretation and attention given to such findings, most of the Marsh study will never be related to future Teacher Corps policy. It is now simply too difficult and tedious a task to be undertaken by the policy makers themselves.

Relation of Marsh study to professional understanding of teacher education. If the findings of the Marsh study for Teacher Corps policy can be considered to be potentially but not realistically useful (because of its form and its lack of interpretation) it should still be considered a major contribution to teacher edu-

cation at large. It is not only an example of the limitations in its approach to measuring the impact of training, but it has also some reliable and provocative data about the performance of student teachers during training and during the first year of employment. Here, the lack of interpretation is fortunate, for the results are so cantankerous, so conflicting with current beliefs about teacher trainee performance, so challenging to current practices and current innovations, that no single interpretation would be believed. Regardless of its many faults as an evaluation study, the Marsh study has generated more evidence that conflicts with current norms than any previous research study into student teacher performance characteristics. Teacher educators should be challenged to make their own interpretations of many of the findings from the Marsh study. The profession can ill afford to disregard the data collected or to forget the challenge that the data make to current professional practices and beliefs.

Discussion

As stated previously, the purpose of this paper was to determine why program evaluations are so often ineffective either for helping determine program policy or for generating knowledge for the profession at large. These six evaluation studies of Teacher Corps demonstrate the variety of approaches evaluators have taken in evaluating a dynamic program and their unfortunate ineffectiveness in providing information for Teacher Corps policy or in generating knowledge for the profession. The rather lengthy review of these studies was needed for two reasons. First, they are relatively unknown. Second, (and most salient to this discussion), it is important to realize that these studies were energetic, highly professional efforts at evaluation. If they have failed as evaluation studies (and they have failed), they have not done so because of a lack of professional expertise, a lack of attention or effort, a lack of creativity or insight or a lack or respect for the complexity of Teacher Corps and the extent of the evaluation challenge. It is hoped that this has been made abundantly clear from the not-so-brief review of these studies.

The following discussion is meant to dwell further upon the

irony of these studies, that is, that while they are creative, insightful professional efforts, they have generated little useful information.

A question that this discussion will try to answer is: Are the inadequacies of these program evaluations due to not following standard evaluation procedures, or are current standard evaluation procedures themselves inadequate for many evaluation purposes? In order to help answer this question, two perspectives of evaluation will be introduced. One perspective is that offered by Alexander W. Astin and Robert J. Panos in their chapter found in the second edition of *Educational Measurement* and titled "The Evaluation of Educational Programs."[37] The Astin and Panos perspective of evaluation will be used to define what is referred to here as "standard evaluation procedures." Their perspective is used rather than others not only because it can be found in a recent handbook of measurement but also because it is consistent with the views expressed by other major figures in the educational program evaluation profession (e.g., Stufflebeam, Popham, Scriven, Provus). In addition, the Astin and Panos perspective is used because it has one of the widest interpretations of the input, process, output model that so characterizes standard evaluation perspectives. The Astin and Panos perspective is used because it is as clear, concise and far ranging a perspective on the input/output model of evaluation as one could expect. It is not a narrow view.

At the same time, although it was meant to synthesize the major characteristics of program evaluation, it is not the only view. There are views of evaluation that do not accept the input, process, output rationale (e.g., MacDonald, Stake, Elliott, Eisner, Hernandez-Nieto, Glass). There are some perspectives on program evaluation that have been offered as alternatives to standard concepts of evaluation procedures. The alternative perspective used in this analysis comes from Malcolm Parlett's and David Hamilton's paper titled "Illuminative Evaluation."[38] Parlett and Hamilton's view is used not because it is the only alternative available,[39] but because it is a clear, concise, and well written alternative. It is also accessible.

Figure 1 introduces the two evaluation perspectives by describing in phrase form their general views about:

- the salient features of educational programs,
- the purposes of program evaluation and
- the nature of the questions that each suggests are appropriate and most crucial to answer in an evaluation study.

Even a cursory look at Figure 1 suggests the contrast between these two perspectives. The standard view of educational programs, for example, is that they are transferable, are located as an instructional method, a text or a course of study and are designed to produce changes in behavior. In the view of Parlett and Hamilton, on the other hand, any educational program is unique, teaching and learning are defined by the participants, and intentions are often left with unintended consequences of unintended processes.

The purpose of program evaluation, for Astin and Panos, is to provide relevant information about causal relationships between inputs, outputs and operations of the program. The purpose is to identify operations that maximize performance, where performance is the achievement of clearly stated objectives. An alternative purpose, as Parlett and Hamilton suggest it, is to understand the process that occurs rather than lessen the sense of uncertainty one has about cause and effect relationships in the educational process.

These differences in the general outlook of what is important to each of the two views of evaluation are reflected by the kinds of questions each would expect answered in an evaluation study. For Astin and Panos, the questions would include: Did the chosen alternative produce the expected outcome? What are the observed relationships between educational operations and educational outputs relative to educational inputs? For Parlett and Hamilton, the questions would include: How has the program operated? What are the relationships between beliefs and practices?

When using these to view or analyze the six Teacher Corps evaluation studies, it is clear that the studies approached their problem as Astin and Panos suggest. All six studies attempted to define the Teacher Corps program in transferable, non-idiosyncratic terms. The RMC I study was so intent upon identifying the objectives of Teacher Corps that that is principally all it did. The Marsh study focused so much upon the causal

Figure 1. Comparison Between Two Evaluation Perspectives

Astin & Panos	Parlett & Hamilton
View of Education Programs	
• a program is an instructional method, single lesson, programmed text, complete course, remedial program, entire school system	• a program is the individuals and group who in turn produce a unique pattern of circumstances, customs, opinions and work styles which suffuse the teaching and learning that are there
• is designed to produce changes in behavior to exposed individuals	• has pervasive operating assumptions
• has some recognized objectives	• is governed by unique teacher characteristics
• is predicated upon belief of the existence of causal relationships between educational objectives and institutional means	• where causal relationships include unintended as well as intended outcomes
• is comprised of outputs, inputs and operations (3 conceptually distinct entities)	• "inputs" include phenomena of crucial educational significance and are difficult to ascertain before instruction (e.g. boredom, interest, concentration, floundering and intellectual dependency)
• a program's objectives are its outputs (e.g. its goals, criteria, outcomes, achievements, dependent variables)	

Purpose of Program Evaluation

- to provide relevant information about causal relationships between inputs, outputs and operations of the program
- to select those educational operations that are most likely to maximize the student's performance on desired outcomes
- to determine different effects of educational operations on educational outputs
- to produce information for educational decision making based upon clearly stated objectives and the achievement of behavioral manifestations of the objectives

- to sharpen discussion, disentangle complexities, isolate the significant from the trivial and to raise the level of debate about the program
- to understand the process occurring within the learning milieu rather than the outcomes
- to provide an informed account of the innovation
- is likely to increase rather than lessen the sense of uncertainty in education
- to move beyond helpless indecision or doctrinal assertion in conduct of instructional affairs

Major Questions

- What was the students' progress or achievement?
- What was the impact of the educational program?

- How has the program operated?
- How has the program been influenced by the various school situations in which it has been applied?

(Continued on page 62)

Figure 1. Comparison Between Two Evaluation Perspectives (cont'd.)

Astin & Panos	Parlett & Hamilton
• Did the chosen alternative produce the expected outcome? • Were the specific intermediate objectives reached? • What were the differences between the students' input and output performances? • How are the outputs best defined and measured? • What are the observed relationships between operations and educational outputs relative to educational inputs?	• What do the participants (and others directly concerned) regard as its advantages and disadvantages? • What is it like to be participating in a scheme, whether as teacher or student? • What may be the program's most significant features, recurring concomitants and critical processes? • What are the relationships between beliefs and practices?

relationships between inputs, outputs and operations that their most interesting, surprising and trustworthy data on student performance went mostly unnoticed and uninterpreted. The Abt study was so concerned with its assumed responsibility for determining specific cause-effect relationships between planned operations and outcomes that the reported results from its surveys and quantification of the case studies fairly contradict or at least disregard most of the information contained in those case studies. The Harvard report, concluding that evaluation was logically a nemesis to a political agency like Teacher Corps, never realized the irony that it was itself an evaluation study with a view of education, a set of purposes, questions and procedures that contradicted their own definition of evaluation. In short, the Harvard study approached its own evaluation task in the manner suggested by Parlett and Hamilton (for example, it emphasized pervading operating assumptions and the milieu in which it was formed) but defined evaluation strictly along the lines suggested by Astin and Panos. One can only conclude from this how pervasive was the sense of standard evaluation to that group.

The point to be made here, however, is not only that these evaluation studies were clearly using the perspective of standard evaluation to define their focus, purpose and major questions, but that they were adhering to this perspective very well. Each study, in its own way, showed creativity and initiative in using these perspectives, that can only be admired. The accomplishment of RMC I in specifying intermediate objectives and long range objectives for a young Teacher Corps, the creativity of the mapping sentence designed by the Apt study to relate causal relationships between input, process and output variables, the extent to which both the Marsh and Corwin studies focused upon student performance as output variables and pursued their *a priori* questions in terms of program operations, all suggest the extent to which the perspective of standard evaluation was not only used, but also was creatively followed in an attempt to evaluate Teacher Corps.

From the Parlett and Hamilton perspective, however, these studies would not be judged so positively. First, the lack of attention to understanding the training processes being used by

different projects points out that the Parlett and Hamilton perspective was not shared by these studies. Operating assumptions and unintended outcomes were seldom considered by these studies as being important features of the Teacher Corps program. For example, although the Abt study, in its case studies, pointed out fundamental unintended consequences of the Teacher Corps projects they described (e.g., the support of the *status quo* in institutional racism), the study never highlighted these results nor addressed them through their interpretations or recommendations. The basic questions of Parlett and Hamilton were never asked. How the programs operated or what the relationships were between beliefs and practices are yet to be described. Clearly, these six Teacher Corps studies approached the challenge of evaluating this dynamic and unusual agency using the basic principles of a standard perspective on evaluation rather than from an "illuminative" perspective.

Figure 2 further compares these two perspectives on evaluation. Figure 2 includes the suggestions for:

- study design features,
- representative variables,
- data gathering procedures,
- data anlysis and
- the nature of the final reports, from the Astin and Panos perspective and from the Parlett and Hamilton perspective.

The contrast between these two approaches should be apparent.

From the perspective of Astin and Panos, a study design would include early definitions of outcomes, exercise of clear control over salient variables and definite statement of the causal relationships between operations and outcomes. From the Parlett and Hamilton perspective, however, an evaluation study would not necessarily be charted in advance, would start with as little control as possible so as to emphasize the day to day reality of the events that are being described and would continually cross-check, re-evaluate and re-design its views of causal relationships. Clearly, the two perspectives differ in what each considers to be an appropriate design for an evaluation study.

Likewise, Figure 2 indicates that the two perspectives differ on:

- potential variables to be considered (e.g., the chosen curriculum vs. the program designers' tacit assumptions),
- appropriate data gathering procedures (e.g., direct measurement of concrete outcomes vs. presence of a wide variety of related events) and,
- corresponding data analyses (e.g., statistical analyses of relationships) vs. observing, inquiring and seeking to explain).

As Figure 2 also shows, even the standard perspective would be designed for program administrators and include information on inputs, outputs and causal interpretations. From an illuminative perspective, on the other hand, a final report would try to present a program as a "recognizable reality" to outsiders with the added criterion that enough evidence be presented (including the critical processes and theoretical principles used) for others to make alternative interpretations.

An analysis of the six Teacher Corps evaluation studies from these two perspectives of evaluation provides additional evidence that the studies not only assumed the basic principles of a standard perspective but also performed credibly from that perspective. The variables chosen in the evaluation studies, for example, were consistent with those suggested by Astin and Panos. These variables included an identification of a chosen curriculum (e.g., CBTE—the Marsh study), a chosen teaching method (e.g., "individualized instruction"—Corwin), the established rules and regulations (e.g., agency guidelines—Corwin and Marsh studies) and treatments or instructional techniques (RMC II and Marsh). Judging by the principles laid out by Astin and Panos, these studies chose the variables "correctly."

An assessment of the choice of variables from the perspective of Parlett and Hamilton, however, would not be so consistently positive. Although the variables identified by the Corwin report would be rated relatively high, using Parlett and Hamilton's perspective (since it did focus upon institutional constraints, status differences and career ambitions and anxieties) the two RMC studies would have to be rated very low since none of the variables suggested by Parlett and Hamilton held much importance. Variables like designer's tacit assumptions, status dif-

Figure 2. Comparison of Design and Procedures of Two Evaluation Perspectives

Astin & Panos	Parlett & Hamilton
Study Design Features	
• early definition of outcomes	• course of study not charted in advance
• exercise control over differential student inputs	• focus on day-to-day reality of settings
• adequately control variables	• progressively clarify and re-define problem areas
• replicated simultaneously in several different social settings	• continual cross-checking of tentative findings
• statement of causal relationship between educational operations and educational outcomes	• systematically reduce the breadth of inquiry to give more concentrated attention to the emerging issues
• existence of two or more independent sets of educational operations for meaningful comparison	
• comparative data based on a comprehensive taxonomy of educational environments	

Representative Variables

- the chosen curriculum
- the chosen teaching method
- design and location of classrooms
- the establishment of certain rules and regulations
- students' achievements, knowledge, skills, aptitude for future learning, values, personality, interpersonal relations
- treatments, learning strategies, teacher style, instructional techniques
- organization of learning experiences, manipulative structuring of learning environment
- program designers' tacit assumptions
- status differences between participants, patterns of attendance, seating, utilization of time and facilities, students' boredom, interest, concentration, floundering, intellectual dependency
- what work has been done, what participants think of their work (the use and value)
- participants' comparisons with past experiences
- career ambitions, anxieties
- intellectual habits
- institutional constraints

Data Gathering Procedures

- incorporate a relatively broad and diverse set of potential objectives
- direct measurement and observation of concrete outcomes
- no data gathering method used exclusively or in isolation
- emphasis on observing and interviewing documentation of day-to-day activities

(Continued on page 68)

Figure 2. Comparison of Design and Procedures of Two Evaluation Perspectives (cont'd.)

Astin & Panos	Parlett & Hamilton
• multiple criteria used to measure achievement of educational objectives • measurement of student inputs and student outcomes (measurement of educational environments much less well developed or certain) • careful planning and judicious selection of variables • objective taxonomic information on several programs	• be present at a wide variety of related events • assemble an information profile from: observation, interviews, questionnaires, tests, documentary and background sources • use open-ended comments • compile work diaries, examples of students' assignments • perspective of how the program was regarded by different people before it began
Data Analyses	
• statistical analyses chosen to produce empirical information on the relationships between outputs, inputs and operations • statistical analyses chosen whose assumptions can be satisfied (e.g. simple covariance often not appropriate)	• three stages of analysis: 1) observe, 2) inquire further and 3) seek to explain • continuous weighing of alternative interpretations in light of new information obtained • not relating different test scores, but accounting for them using the study's

findings
- cross-check most important findings
- use of outside consultants to challenge preliminary interpretations,
- play devil's advocate, check open-ended material and develop own interpretations

Nature of Final Report

- designed for program administrators and instructors
- designed for formulation and refinement of educational theory
- include information on the student outputs
- include information on the student inputs
- all plausible rival interpretations considered before causal interpretations given
- replace existing folklore with research information
- present program so that it can be seen as a recognizable reality by others outside the innovation
- may increase rather than lessen the sense of uncertainty
- document critical research processes
- theoretical principles and methodological ground rules discussed and made explicit
- criteria for selecting or rejecting areas of investigation spelled out
- present evidence in such a way that others can judge its quality and make alternative interpretations

ferences between individuals, patterns of utilizing time, student boredom or student intellectual dependency were of little interest to these studies. Thus, from the Parlett and Hamilton perspective, the variables chosen for these studies would generally be inappropriate and consistently be incomplete.

There are also contrasts in the data gathering procedures suggested by these two perspectives on evaluation. Although both, for example, suggest a wide range of data gathering procedures, Astin and Panos emphasize the objectivity of these observations, the measurement of inputs, outputs and (where possible) procedures. Again, along the lines suggested by Astin and Panos, these Teacher Corps evaluations were uncommonly professional. Most incorporated a broad range of potential objectives, directed their measurement to "concrete" outcomes, measured student input and output, used careful planning and judicious selection of variables and, to the best of their capabilities, provided objective taxonomic information on several programs. The recommendations of Parlett and Hamilton that differed from the standard recommendations for data gathering, however, were not incorporated into these studies. Documenting day-to-day activities, for example, or being present at a wide variety of related events were not part of the procedures used to gather data. Thus, although there are commonalities between the data gathering procedures suggested by Astin and Panos and those suggested by Parlett and Hamilton, one would still expect a very positive judgment using the criteria suggested by Astin and Panos and a more negative judgment using the criteria suggested by Parlett and Hamilton.

The data analyses used in these evaluation studies must be considered as practical compromises with the two principles stated by Astin and Panos. That is, the obvious intentions of these studies were:

- to produce empirical (statistical) information on the relationships between outputs, inputs and operations and
- to chosse statistical procedures whose assumptions could be satisfied.

In practice, the compromise with these principles included using available statistical procedures to produce correlations

between the chosen variables and to apologize later for the extent to which basic assumptions of these procedures were violated. Multiple linear regression and various orthogonal transformations were used, for example, by the Corwin and Marsh studies, to reduce the number of factors and arrive at more interpretable results about their input, output and process variables. Although Corwin seemed to be quite successful in applying these procedures, he also made clear the extent to which basic assumptions such as the additivity and constancy of treatment effects and the lack of interference between experimental units just could not be satisifed given the realities of Teacher Corps and its projects. In fact, the most relevant information reported in Corwin (Chapter Eleven) was reported as conjecture, interpretation and hypothesis. In short, using the criteria of Astin and Panos, the data analytic procedures chosen in these studies would have left much to be desired.

Using the criteria of Parlett and Hamilton, it is difficult to arrive at a simple assessment of the data analyses used. Each study eventually realized that it was encountering more than it had anticipated. From the Harvard report, which came to a conclusion that contradicted its original premise (i.e., that standard evaluation was responsible for effective conduct of a national agency), to the RMC, Abt, and Corwin studies, which included stages of further inquiry, some weighing of alternative interpretations and trying to account for the "findings" with other information such as case study information or informal contacts, the recognition of possible pluralistic interpretations was implicit. From the Parlett and Hamilton perspective, however, it was unfortunate that these interpretive difficulties were not anticipated nor planned for in the study designs. The lack of alternative interpretations or the minimal efforts given to continually alter interpretations as the studies progressed would suggest that the basic principles of analysis offered by Parlett and Hamilton were not followed. Nevertheless, in reading the conclusions of these studies, one would be led to say that these principles of analysis were more tacitly followed as the limitations of statistical results became more apparent. All studies, for example, reported a learned appreciation for the complexities, ambiguities and contradictions represented

within those Teacher Corps contexts that were being evaluated.

The nature of the final reports of these six studies consistently followed the principles of Astin and Panos. They included information on student inputs and outputs and clearly were directed at replacing "folklore" with "research information." The attempts to reduce conclusions to specific variables reflected an intent to write to program administrators. If one were to seriously assess the extent to which these reports would have been clearly understood by program administrators, however, these reports would probably receive very low grades. Likewise, as has been suggested previously, except for the Corwin study which used a perspective of organizational theory, there was little refinement of formulation of educational theory.

From the viewpoint of Parlett and Hamilton, the final reports of these six evaluation studies were utter failures. In no way was there an attempt to present the program as a recognizable reality to others. Evidence was not presented for alternative interpretations. The point of these reports was definitely to lessen (rather than increase) the sense of uncertainty. And, except for the Corwin and Abt fascinations with organizational theory, there were few theoretical principles and methodological ground rules made explicit. In short, although the final reports may have been found wanting in some aspects from the Astin and Panos perspective, they would have been totally unacceptable from the view of Parlett and Hamilton.

Ironically, the study design offered unique problems. Each study, as Astin and Panos suggest, did set out to accomplish an "early" definition of outcomes, did devote obvious energies to stating anticipated causal relationships between these variables, and did try to make comparisons of comprehensive taxonomies of "environments." Each also, however, reported the extent to which they simply could not control the variables they had identified, such as student input, training procedures, project setting and project operating characteristics. This lack of control by the evaluators of a complex program such as Teacher Corps is not surprising. Nevertheless, it does point out the difficulties the six studies had in applying certain standard study design features to the evaluation of Teacher Corps projects.

It is also clear, however, that these studies did not use a Parlett and Hamilton perspective in designing their studies. The focus was definitely not on the day-to-day reality of the settings; problem areas were not dwelt upon for clarification and redefinition; continual cross-checking was used only in special case studies rather than in the "final" analyses, interpretations and recommendations. From a Parlett and Hamilton perspective, the design of these studies was not aimed in the right direction. From an Astin and Panos perspective, the aim of the design was there but they fell short of the mark, mostly because of conditions that could not be controlled.

Conclusion

These six evaluation studies performed credibly from the perspective of standard evaluation. The questions that were asked, the variables that were chosen, the data gathering procedures used and the nature of the final reports not only followed many of the criteria suggested by Astin and Panos but did so creatively. Even some of the failures from the Astin and Panos perspective (lack of control over the variables, or inability to satisfy the assumptions necessary for the statistical analyses chosen) suggest the extent to which the basic principles of the standard perspective of evaluation were applied. There can be no doubt that these six large scale evaluations of Teacher Corps clearly applied the standard perspective on evaluation as best they could, given the complexities of Teacher Corps concepts, goals, guidelines and activities.

In addition to the conclusion that these studies performed credibly according to principles of standard evaluation, there is unmistakable evidence that these studies were not (and could not be) useful to either Teacher Corps, to high level agency decision makers or to professional understanding. Major testimony from these studies is not only to the perseverance, hard work and (at times) creativity with which these principles were applied, but also to the general lack of utility of these principles for generating knowledge about either the program, or its effects, or for giving lessons about its successes and failures.

From the evidence of these six evaluation studies of Teacher Corps, one is forced to conclude that the failure of these studies

is not a failure to follow suggested procedures. The principles stated by Astin and Panos are inadequate for their own purposes when evaluating federal agencies such as Teacher Corps.

Where does this leave us? Certainly there is no proof here that an alternative concept of evaluation will work. Nevertheless, the discussion using the Parlett and Hamilton perspective shows that there are perspectives of evaluation that are (as they argue) alternatives to standard perspectives. The conclusion from this review of evaluation studies is that it is necessary to try out federal program evaluation from perspectives other than those represented by Astin and Panos. Parlett and Hamilton's "illuminative evaluation" is one candidate among many. Other possibilities include MacDonald's notion of "democratic evaluation," Stake's concept of "responsive evaluation," Hernandez-Nieto's suggestion of "process-evaluation" or Bronfenbrenner's concept of "ecological experiments."[40] Considering the amount of professional effort and creativity expended in previous attempts to evaluate Teacher Corps and the corresponding absence of information about the workings of the agency, its local projects, the training procedures intended, altered and used or the effectiveness of this training, there is no choice but to encourage the use of alternative methodologies in evaluating complex federal programs.

Recommendations

Recommendations from the above conclusion are so obvious that they need hardly be mentioned. These include:

- the need of the Office of Planning, Budget and Evaluation (OPBE) for more flexible evaluation guidelines and principles,
- the need for more systematic accounts of alternatives to the current standard perspective of educational evaluation and
- the need to provide more training and experience in those procedures necessary to perform evaluations from alternative perspectives.

In addition to the recommendation that different perspectives of evaluation be practiced and supported in evaluating federal programs, it is recommended that the following suggestions be considered. The point being raised by these

suggestions is that too much is taken for granted about federal program evaluation. These suggestions are made to encourage further discussion about the possibilities for:

- providing more information that can be useful in determining policy, or
- further illuminating professional and public understanding of educational processes that are being described and assessed.

Suggestion: Have professionals evaluate the agency who are themselves responsible for determining or carrying out policy on similar programs.

For example, if Teacher Corps is to be evaluated again, why not include:

- teacher representatives,
- school of education personnel,
- state department of public instruction personnel, or maybe,
- an international team from similar backgrounds and concerns,

as part of the evaluation effort? Part of the rationale for this suggestion is that the most that can be done now is to offer information that allows policy makers to experience vicariously the consequence and meanings of certain policies and activities in order that better (more informed) policy can be made in the future. There are, of course, limits to what evaluators can communicate to "decision makers." It should also be realized that the process evaluating an agency as complex as Teacher Corps has in itself, a powerful effect on the evaluators. The impact of the evaluation process upon the evaluators can be tapped for professional purposes rather than dissipated to private evaluation firms.

Suggestion: As much effort should be expended in making the reports clear, interesting and useful as there is in collecting, analyzing or interpreting the results.

Why are so many evaluation reports so difficult to understand and boring as well as useless? One reason may be the lack of time and effort given to communicating or demonstrating the "findings." Not only is it necessary to identify the audiences and write to them; it is also necessary to choose the most direct way to make the messages understood. One may expect, for

example, different kinds of reports for different audiences. The possibilities for the nature of reports may be endless but these possibilities are determined by two factors. One is the imagination of the evaluators (along with their knowledge of the audiences); another is the amount of time, effort and expertise that is made available. Successful communication of results must be the major criterion of the value of evaluation.

Suggestion: Continually (or at least periodically) include client policy makers in the interpretive procedures.

For an agency like Teacher Corps there are at least three sets of policy makers. One is the legislature which determines the funding levels and mandates the purposes; another is the parent agency (Office of Education) to which the agency reports; the third set is the agency personnel themselves, including the director, staff and local (and future) project directors. Each policy level has its own set of questions and decisions that can change dramatically over time. The suggestion here is not only to make sure that changes in policy priority and in evidence and interpretation are monitored by both evaluators and their clients, but to realize that each must have a realistic perspective both of the concerns of the others and the limits of the findings.

Suggestion: Conduct federal agency evaluations as interpretive efforts rather than research results.

One of the major conclusions of this review was that most of what is given as "conclusions" is little related to "fact." The results, the instruments, the procedures and data analytic methods to arrive at the results, although occasionally interesting (and perhaps "the best there is") are little related to the most useful, helpful and insightful statements. Remember, these are evaluation studies. More precisely, they are efforts to demonstrate the success or failure of certain intervention strategies. Conclusions from these studies do not come directly from the data and data analyses. Thus, the expertise needed to interpret the data is of a far different sort than that needed to form conclusions in tightly controlled educational research. What is suggested here is an approach to the study in the following vein:

- this is what we wanted to know and why we wanted to know it,
- this is how we tried to understand what occurred, and

- these are our interpretations and recommendations from what we have learned.

In short, there is no such thing as unambiguous data in these reports. Wisdom, past experience, understanding, creativity and the professional values held by the participants in the evaluation will be the governing factor in the usefulness, "validity" and excitement of the recommendations. Technical expertise of a researcher is simply not enough. Thus, it becomes very important to choose a team of seers and idea people (preferably from different perspectives) to design, gather, interpret, report and recommend.

Suggestion: Consider the needs of the profession for unique information or unusual perspectives and interpretations of their practices along with the needs of policy makers.

Evaluation studies often provide unique opportunities to the profession at large. They offer the profession opportunities to observe natural field settings, especially to observe unusual personal, group or institutional efforts being expended in these settings. The potential for field information that can both shape and re-make professional perspectives is too great not to be taken into account when considering the design of an agency evaluation. Evaluation offers challenges to the professional perspectives normally used to describe natural settings and also challenges ideas as to how lessons are or can be learned from these settings. Another reality of agency evaluation is that *every* evaluation report is an historical document (as well as a policy statement) as soon as it is written. Keeping this in mind, it makes sense to record events, processes, intentions and outcomes (intended and unintended) as historical data describing a particular case rather than as a timeless proof of a universal principle. No matter how information is to be collected, an evaluation study can be designed so that experiences and lessons from the agency can be shared with the profession as well as with those directly responsible for agency activities.

Suggestion: Support evaluation methods that can capture growth, dynamic interplay, liveliness and impact (as well as the absence of these qualities).

The ways to capture and interpret dynamic settings must be addressed creatively with few professional restrictions upon

what is allowed as "evidence." Currently what is missing from evaluation studies is any sense of the liveliness, growth, periodic failures and subsequent redirections that make up most of the existence of an interventive project. Without ways to capture these features of intervention strategies, there is no reason for participants, interested bystanders, or decision makers to accept the description as real or the interpretation as valid. For example, one common characteristic shared by dynamic features is time, yet time is a nonvariable in most evaluation studies (even for those in which "process" is described).

Current events, past failures and immediate needs for more understanding of the information that has been collected leave us with no choice but to search for an alternative evaluation paradigm. As this discussion has emphasized, that alternative must include a greater understanding of the "black box" of educational processes, must approach "treatment" itself as a problematic, dynamic variable, needs more than *a priori* definitions and categories of intentions (for both process and outcome), allows for alternative assumptions about what is to be termed "success" and "failure," and focuses not only upon what are common sense variables but also upon the extent to which these variables can and cannot be manipulated in specific settings. In short, an alternative is needed that is appropriate for describing, illuminating and interpreting the dynamics of a particular program.

This review of past studies of Teacher Corps has suggested that "better performance" of current standard evaluation procedures cannot be expected to produce better results. It must also be realized, however, that there is no conclusive evidence that an alternative set of procedures will be more useful either to policy determination or to professional understanding. The Parlett-Hamilton perspective has only suggested that perspectives and accompanying methodologies exist that offer different criteria for excellence and different possibilities for success. We will never know more than that alternatives exist until those alternatives, too, can be subjected to the difficult challenges of federal agency evaluaton. Considering the current need for more appropriate methodologies, the possibilities for learning much more about what would constitute successful programs

and the existence of alternative evaluation procedures, our failure to employ these alternative designs, at least experimentally, is gross professional negligence. In conclusion, this critical analysis of past Teacher Corp evaluations suggests that we have used an extremely arid approach to measure our Springs.

References

An Assessment of Teacher Corps. 2 volumes. Bethesda, Md.: RMC Incorporated (Robert Crosby, Director), August, 1970. Grant no. OEC-0-099018-4647.

Apple, Michael. "The Process and Ideology of Valuing in Educational Settings." In Apple, Subkoviak and Lufler. *Educational Evaluation: Analysis and Responsibility.* Berkeley: McCutchan Publishing Corporation, 1974.

Astin, Alexander and Panos, Robert. "The Evaluation of Educational Programs." In Thorndike, Robert L. *Educational Measurement*, Second Edition. Washington, D.C.: American Council on Education, 1971. pp. 733-751.

Bloom, B., Hastings, J. and Modans, G. *Handbook on Formative and Summative Evaluation of Student Learning.* New York: McGraw-Hill, 1971.

Bronfenbrenner, Uri. "The Experimental Ecology of Education." *Educational Researcher 5* (9) October 1976, pp. 5-15.

Brugelman, Hans. "Towards Checks and Balances in Educational Evaluation: On the Use of Social Control in Research Design." In McDonald, B. and Walker, R. *Innovation, Evaluation, Research and the Problems of Control.* East Anglia, England: SAFARI PROJECT Interim Papers. Centre for Applied Research in Education, 1974.

Bussis, A., Chittenden, E. A. and Amarel, M. *Methodology in Educational Evaluation and Research.* Unpublished manuscript. Princeton, NJ: Educational Testing Service, 1973.

Campbell, Donald T. "Degrees of Freedom and the Case Study." *Comparative Political Studies. 8* (2), July 1975, pp. 178-193.

Campbell, Donald. "Reforms As Experiments." In Struening, E. and Guttentag, M. *Handbook of Evaluation Research, Volume 1.* Beverly Hills: Sage Publications, Inc., 1975.

Carini, P. F. *Observation and Description: An Alternative Methodology for the Investigation of Human Phenomena.* Grand Forks, ND: University of North Dakota, North Dakota Study Group on Evaluation, Center for Teaching and Learning, 1975.

Chapter IV: *The Role of Evaluation in the Bureau of Educational Development.* "The Teacher Corps: A Case Study in Evaluation." Boston: Center for Educational Policy Research, Harvard University, December, 1971.

Cicirelli, V., Evans, J., Schiller, J. "The Impact of Head Start, A Reply to the Report Analysis." *Harv. Educ. Res. 40,* 1970, 105-129.

Coleman, James. "Recent Trends in School Integration." *Educational Researcher, 4* (7), 1975, pp. 3-12.

Coleman, James, et al. *Equality of Educational Opportunity.* Washington, D.C.: U.S. Government Printing Office, 1966.

Cort, H. Russel, Jr., and O'Keefe, Ruth Ann. "Teacher Corps—Two Years of Progress and Plans for the Future." Reports prepared under a contract from the Office of Education, Washington, D.C. October 15, 1968. Monograph.

Corwin, Ronald G. *Reform and Organizational Survival: The Teachers Corps as an Instrument of Educational Change.* New York: Wiley-Interscience, 1973.

Corwin, Ronald. *What Difference Has the Teacher Corps Made? Longitudinal Studies of the Teachers Corps Program.* Report submitted to the Ford Foundation through the NEA. September 1973.

cummings, e. e. *100 Selected Poems.* New York: Grove Press, 1923.

Davis, Don and Amershek, Kathleen. "Student Teaching." In Ebel, Robert L. *Encyclopedia of Educational Research*, 4th edition. New York: MacMillan and Co., 1969.

Dusalt, Gilles. *A Theory of Supervision in Teacher Education.* New York: Teacher College Press, 1970.

Eisner, E. "Emerging Models for Educational Evaluation." *School Review. 80*, 1971, pp. 573-590.

Eisner, E. "The Perceptive Eye: Toward the Reformation of Educational Evaluation." Stanford Evaluation Consortium, Department of Education, Stanford University. Occasional Paper, December, 1975.

Federal Register. 40 (243) Wednesday, December 17, 1975. pp. 17-43.

Fox, G. T. "Personal Notes and Reactions to Marsh Report" and "Revision of Variables of Marsh Report." Mimeo, University of Wisconsin-Madison. January, 1974.

Fox, G. T. (ed.) *Federal Role in School Reform From Sociological And Educational Perspectives.* Madison, Wisconsin: University of Wisconsin, 1974. Published through a grant from Teachers Corps Projects, Office of Education.

Fox, G. T. Grant, C., Popkewitz, T., Romberg, T., Tabachnick, B. and Wehlage, G. *1975 CMTI Impact Study Technical Reports Numbers 1-21.* Madison, Wisconsin: University of Wisconsin-Madison. September 1976. Prepared as part of Teachers Corps, Office of Education contract no. 300-75-0100 (University of Nebraska at Omaha).

Full-Scale Implementation of a Process Evaluation System for Programs of NCIES. Report UR-172, Volume 1, and Part G. Volume 2. Bethesda, Md.: RMC Incorporated (Robert Crosby, Director), November, 1972.

Glass, Gene. "A Paradox About Excellence of Schools and the People in Them" *Educational Researcher.* March 1975, pp. 9-13.

Glass, Gene (ed.) *Evaluation Studies Review Annual. Volume 1.* Beverly Hills: Sage Publications, 1976.

Gottman, J. M. and Clasen, R. E. *Evaluation in Education. A Practitioner's Guide.* Itaska, Ill.: F. E. Peacock Publishers, Inc., 1972.

Guba, E. and Stufflebeam, D. *Evaluation. The Process of Stimulating, Aiding and Abetting Insightful Action.* Monograph Series In Reading Education No. 1. Bloomington: Indiana University, 1970.

Haberman, Martin. "Twenty-three Reasons Universities Cannot Educate Teachers." *The Journal of Teacher Education.* Volume XXII, No. 2, Summer,

1971.

Hernandez-Nieto, Rafael. *Process-Development: An Alternative Approach for Curriculum Research in Mathematics.* Madison, Wisconsin: Wisconsin Research and Development Center for Cognitive Learning. Working Paper No. 197. November, 1976.

House, Ernest. "Justice in Evaluation." In Glass, G. *Evaluation Studies Review Annual, Volume I.* Beverly Hills: Sage Publications, 1976.

Innovation and Change: A Study of Strategies in Selected Projects Supported by NCIES. 4 volumes. Cambridge, Mass.: Abt Associates, December, 1972. Grant no. OEC-9-71-3714.

Jackson, Gregg. "Reanalysis of Coleman's Recent Trends in School Integration." *Educational Researcher.* 4(10) 1975, pp. 21-25.

Karier, Clarence. *Ideology and Evaluation: Analysis and Responsibility.* Berkeley: McCutchen Publishing Corporation, 1974.

MacDonald, Barry. "Evaluation and the Control of Education." In MacDonald, B. and Walker, R. *Innovation, Evaluation, Research and the Problem of Control.* East Anglia, England: SAFARI PROJECT Interim Papers, Centre for Applied Research in Education, 1974.

MacDonald, Barry and Parlett, Malcolm. "Re-thinking Evaluation: Notes from the Cambridge Conference." *Cambridge Journal of Education, 3* (2), Easter Term, 1973. pp. 74-82.

Marsh, David and Lyons, Margaret. *A Study of Teacher Training at Sixth-Cycle Teacher Corp Projects. Volume I. Methodology and Findings.* July, 1974; *A Study of the Effectiveness of Sixth-Cycle Teacher Corps Graduates.* Phase II. October, 1974; *Teacher Corps Graduates in Highly Innovative Roles in Education: A Follow-Up Study.* Phase II. February, 1975. Pacific T and TA Corporation. U.S. Department of Health, Education and Welfare, Office of Education, Office of Planning, Budgeting and Evaluation. Contract no. OEC-0-73-5174.

Novack, Mark and Gilbert, Sid. "Paradigms and Policy." Paper delivered at 1975 AERA Convention, April, 1975.

Parlett, M. and Hamilton, D. "Evaluation as Illumination: A New Approach to the Study of Innovatory Programs." In Glass, Gene. *Evaluation Studies Review Annual, Volume I.* Beverly Hills: Sage Publications, 1976. pp. 140-157.

Popkewitz, Thomas. "Legitimatizing Failure: Teachers Corps as a Reform." *School Review 84* (1).

Provus, M. *Discrepancy Evaluation.* Berkeley: McCutchen Publishing Corporation, 1971.

Ross, Lee and Cronbach, Lee (ed.) *"Handbook of Evaluation Research,* Essay Review." *Educational Researcher 5* (10) November, 1976, pp. 9-19.

Scriven, Michael. "The Methodology of Evaluation." In *Perspectives of Curriculum Evaluation.* AERA Monograph Series on Curriculum Evaluation. No. 1. Chicago: Rand McNally, 1967.

Scriven, Michael. "Pros and Cons About Goal Free Evaluation." *Evaluation Comment 3* (4) 1972, pp. 1-14.

Scriven, Michael. "Maximizing the Power of Causal Investigations: The Modus Operandi Method." In Glass, G. *Evaluation Studies Review Annual, Volume I.* Beverly Hills: Sage Publications, 1976.

Stake, Robert. "The Countenance of Educational Evaluation." *Teachers College Record. 68*, 1967, pp. 523-540.

Stake, Robert. "An Approach to the Evaluation of Instructional Programs (Portrayal vs. Analysis)." Paper presented at 1972 AERA Convention. Chicago, April, 1972.

Stufflebeam, D., Foley, W., Gephart, Guba, E., Hammond, R., Merriman, H. and Provus, M. *Educational Evaluation and Decision Making.* Bloomington, Ind.: Phi Delta Kappa National Study Committee on Education. 1971.

Suchman, E. *Evaluation Research: Principles and Practice in Public Service and Social Action Programs.* New York: Russell Sage Foundation 1967.

Teachers Corps: Past or Prologue? A Report with Recommendations, by the National Advisory Council on Education Professions Development. Walter Tice, Chairman: George Arnstein, Executive Director. Washington, D.C. July, 1975.

Weiss, Carol. "The Politicization of Evaluation Research." *Journal of Social Issues. 6* (4), 1970, pp. 57-68.

Young, Michael (ed.) *Knowledge and Control.* London: Collier-MacMillan, 1971.

Notes

[1]This is a stanza from a poem, no. 23, "voices to voices, lips to lips", in e.e. cummings (1923).

[2]For example, see Apple (1974), Karier (1974), Novack and Gilbert (1975), Ross and Cronbach (1976), Stake (1967) and Young (1971), who question the ideological or theoretical efficacy of standard evaluation procedures.

[3]For example, Jackson (1975) critically analyzes Coleman data on school integration.

[4]*Teacher Corps: Past or Prologue?*, a report by the National Advisory Council on Education Professions Development, George Arnstein, executive director (1975) offers a concise introduction to Teacher Corps.

[5]The Marsh and Lyons study (1974) is one example.

[6]For example, the Corwin study, *Reform and Organizational Survival* (1973) was funded jointly by NEA and the Ford Foundation.

[7]For example, the Cort and O'Keefe study (1968) was funded by Teacher Corps.

[8]Agencies such as Abt Associates, Incorporated, Resource Management Corporation and Pacific Training and Technical Assistance Corporation have taken part in evaluations of Teacher Corps.

[9]See *Teacher Corps: Past or Prologue?* (op cit).

[10]As evidenced in the *Federal Register*, December 17, 1975, Teacher Corps projects became "demonstration" projects.

[11]Davis and Amershek (1969), Haberman (1971) and Dusalt (1970) are teacher educators who were identifying the need for a greater knowledge base and more empirical research into teacher education at the same time that these evaluation studies were being conducted.

[12]Astin and Panos (1971)

[13]Parlett and Hamilton (mimeo in 1972), published in 1976).

[14]For a smooth and continuous reading, footnotes will only be made in these descriptions of evaluation studies for direct quotes.

[15]In 1975, a technical assistance project was funded for providing Teacher Corps with a Management Information Ssytem (MIS).

[16]*Chapter IV: The Role of Evaluation in the Bureau of Educational Development*, p. 109.

[17]*Chapter IV: The Role of Evaluation in the Bureau of Educational Development, p. 155.*

[18]*Chapter IV: The Role of Evaluation in the Bureau of Educational Development*, p. 156.

[19]*Full-Scale Implementation of a Process Evaluation System for Programs of NCIES*, p. 40.

[20]Ibid., p. 40.

[21]*Innovation and Change*, Volume I, p. 1.

[22]*Inovation and Change*, Volume II, p. 32.

[23]Corwin (1973), p. 11

[24]Corwin (1973), p. 116

[25]Corwin (1973), p. 116-117.

[26]Corwin (1973), pp. 116-117.

[27]Since the Marsh study, the closest effort to generating information about teacher training practices of Teacher Corps projects was the 1975 CMTI Impact Study (Fox, Grant, Popkewitz, Romberg, Tabachnick and Wehlage).

[28]Marsh and Lyons, *Follow-Up Study*, p. 15.

[29]Marsh and Lyons, *Follow-Up Study*, p. 16.

[30]Ibid., for example, pp. 16, 17, 18.

[31]Ibid., p. 18.

[32]See Fox (mimeo) (January 1974).

[33]It reflected, for example, both the views of Scriven (1967) and Stufflebeam *et al* (1971) that the major responsibility of program evaluation is to attempt to determine causal relationships between variables.

[34]Marsh and Lyons, Volume 1, p. 166.

[35]Ibid., p. 16

[36]Marsh and Lyons, *Follow-up Study*, p. 46.

[37]See Astin and Panos. (1971)

[38]See Parlett and Hamilton (1976).

[39]See Carini (1975), Eisner (1975), MacDonald (1974), Stake (1972).

[40]These views were explicated in MacDonald (1974), Stake (1972), Hernandez-Nieto (1976), Bronfenbrenner (1976).

Teacher Corps Evaluation Design

Excerpts from Task Forces Report and from Request for Proposals

This chapter consists of two sections; both of them are edited exerpts taken from previously distributed Teacher Corps documents. The first section contains excerpts from the Evaluation Task Force report[1] of a series of meetings held in the Summer of 1977. The second section is taken from the Request for Proposals (RFP) to conduct the Teacher Corps Evaluations. The RFP has been changed from outline to narrative form for easier reading. Some of the repetition has been removed. Other sections have been rearranged for the convenience of the reader. Those who wish to examine the RFP, as it was distributed for procurement purposes, should consult that document.

EXCERPTS FROM TASK FORCES REPORT

The Task Force Program Evaluation was charged with developing an evaluation design appropriate for Teacher Corps Program 78 efforts. The proposed design described in this report includes 1) the purposes and components of the evaluation, 2) several substantive frameworks which could guide development of the evaluation, and 3) a collaboration and governance structure for the evaluation.

Proposed Design

Based on an analysis of previous Teacher Corps evaluations, the following desirable evaluation design characteristics were identified:.

---

[1]Full reports of all three task forces may be found in Teacher Corps Task Forces, Document I—*Reports on Demonstration, Program Evaluation and Technical Resources*, Developmental Training Activities, Teacher Corps, Center for Urban Education, University of Nebraska, 1977.

- Evaluation efforts should be jointly undertaken by Teacher Corps/Washington and the Office of Planning, Budget, and Evaluation.
- Evaluation should be an on-going process, rather than a single-shot effort.
- Evaluation should be responsive to new evaluation needs as they emerge.
- Evaluation should provide for on-going data collection which emphasizes documentation and qualitative as well as quantitative data.
- The evaluation effort should provide immediate feedback to Teacher Corps/Washington and to local projects so that they can effectively utilize the evaluation data.
- A democratic approach should be taken to the evaluation's design, data gathering, data processing, and interpretation.
- The evaluation effort should facilitate a high level of utilization by Teacher Corps/Washington and local projects.
- The evaluation efforts should respect confidentiality.
- Before their wide dissemination, the evaluation findings should be shared with and critiqued by appropriate professionals.

These characteristics are incorporated in the evaluation design.

Components of Design

The evaluation design has four components, as follows:

- *A Locally Controlled Project Documentation and Evaluation Component* to aid local program development.
- *A National Process-Outcome Assessment* featuring cross-site analysis of process and outcomes, and conducted by an independent evaluator.
- A series of theory-based *Special Project Studies* of project dimensions judged to be most critical to the success of Teacher Corps projects.
- A *Summative Synthesis Evaluation*, which is a secondary analysis of quantitative and qualitative data and which addresses broad policy questions related to Teacher Corps.[2]

Task Force members believe that these four closely related evaluation components can collectively form one overall evaluation effort. It recommends that the National Process-Outcome

[2]Not in Request for Proposal 1978-90.

Assessment be funded and/or monitored by the Office of Planning, Budget, and Evaluation, while the other three evaluation components be funded directly by Teacher Corps/Washington. A collaborative governance mechanism (to be described later in this report) will relate these various evaluation components to each other and to the overall evaluation approach.

A Local Project Documentation and Evaluation Component.

The primary purpose of the local program documentation and evaluation component is to aid local program development. This effort is mandated in the Teacher Corps program guidelines for Program 78 projects. Under those guidelines, each project must

- document all significant factors which influence project experience and results,
- review progress in accomplishing the objectives developed in the original proposal,
- revise those objectives, if necessary.

The Evaluation Task Force proposes that Teacher Corps/Washington aid local projects' documentation and evaluation efforts in several ways. First, Teacher Corps/Washington should support the development of a quality documentation framework, which local projects can then adapt to their needs. Second, training in the use of this framework should be provided to local project evaluators/documentors. The improvement of the locally proposed frameworks should also be assisted by Teacher Corps/Washington. Finally, Teacher Corps/Washington should monitor local documentation and evaluation efforts to ensure quality performance.

A National Process-Outcome Assessment.

The second component of the overall evaluation design is the National Process-Outcome Assessment.[3] The goals of this effort are as follows:

- To describe local project impact in terms of the nine major outcomes of Teacher Corps programs, using common cross-site measures in a time-series perspective.

[3]Refer to current RFP 78-90.

- To describe and document local project context, structures, and processes by applying a set of common measures in a time-series perspective.
- Within financial constraints, to obtain an in-depth assessment of process and outcome indicators at a sample of projects and to incorporate these insights within the broader analysis of process and impact described in Goal 4.
- To examine relationships among and between Goals 1, 2, and 3, with emphasis on providing policy and operational implications for Teacher Corps/Washington and local projects.
- To document and assess the process and impact of technical assistance and other Teacher Corps/Washington interaction with the field and the relationship of these efforts to successful or unsuccessful project implementation.
- To demonstrate a useful new approach to evaluation which emphasizes: thorough documentation, as well as evaluation, of local projects; more effective collaboration between Teacher Corps/Washington, and local projects, OPBE, and outside evaluators; and more effective utilization of documentation, evaluation, and data bank efforts as a vehicle for improving national Teacher Corps and local project efforts.

The Task Force recommends that this assessment be conducted over an eight-year period by an independent evaluator funded through a competitive RFP.[4] This strategy would allow the independent evaluator to examine Program 78 projects throughout the five years of their existence and, further, to examine the pattern of institutionalization of educational improvements following the termination of projects.

The Task Force also recommends that baseline data be collected on the four basic outcomes prescribed in the Teacher Corps *Guidelines*: namely, outcomes related to school climate, education personnel development systems, institutionalization, and adoption/adaptation of project-generated educational improvements by other educational institutions or agencies. The Task Force assumes that baseline data would focus on rather rapid and simple assessments of the four basic outcomes as determined by scales, inventories, or survey instruments. Several related project outcomes might also be assessed; these include the capacity of cooperating institutions to continue to

[4]Contractual/budget restraints limited the RFP 78-90 to five years.

change after their involvement with Teacher Corps. The Task Force recommends that a balance be struck between the need to define the basic outcomes of Teacher Corps in the near future (so that baseline data can be collected appropriately) and the need to allow the meaning of these outcomes to evolve over time.

Special Project Studies.

A set of extremely important special studies will complement the local program documentation and evaluation and the National Process-Outcome Assessment. These complementary studies will provide a theoretical understanding of important issues related to project implementation and institutionalization, as well as those related to federal relationships with local projects. The emphases for these studies will be on alternative conceptual frameworks and data collection processes, and on the utilization of theory from related disciplines to shed light on Teacher Corps processes and results. For example, the following issues might be addressed in special project studies:

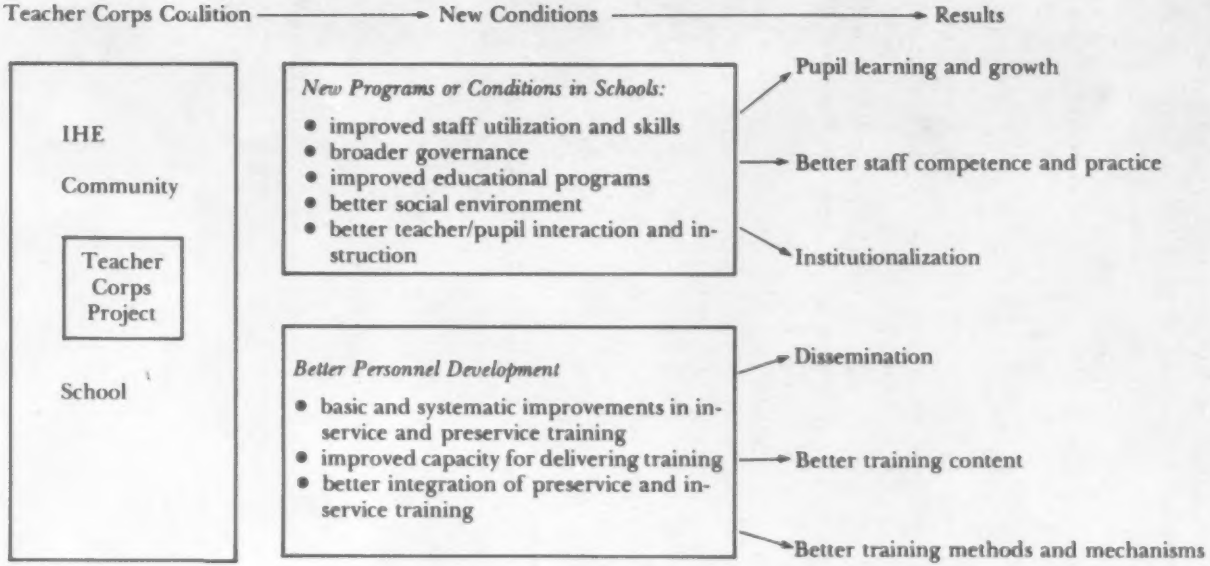
- How does a beginning teacher get socialized into Title I schools? How can one determine which teachers will succeed in Title I schools?
- How do the governance structure and processes within projects relate to successful or less successful accomplishment of the demonstration goal of Teacher Corps?
- How do various modes of Teacher Corps technical assistance relate to project success?

In order to address such issues, the studies might, for example, utilize longitudinal case studies involving interns, teachers, and other trainees over an eight-year span. The purpose of these individual case studies would be to understand project dynamics through the experiences of individual participants. They would also help their audiences understand changes and trends in the participants' own professional careers and viewpoints.

Substantive Frameworks

The evaluation Task Force has identified three approaches to developing substantive frameworks. The first framework is

Figure 3. The Generic Purposes of Teacher Corps



based on the generic purposes of Teacher Corps and is presented in Figure 3. A second approach to developing appropriate substantive frameworks is based on the idea that Teacher Corps/Washington needs a variety of alternative substantive frameworks which could serve specialized purposes within the evaluation design. For example, one substantive framework viewed Teacher Corps as a set of roles (such as institution builder, teacher trainer, institutionalizer, etc.). These roles suggest interesting dimensions along which to document local Teacher Corps dynamics. A third approach to identifying a substantive framework is based on "unpacking" the Program 78 *Guidelines*. This approach would help to clarify the meaning of school climate and educational personnel development systems, for example.

Governance and Collaboration

The Evaluation Task Force feels that Teacher Corps/Washington and the Office of Planning, Budget, and Evaluation should jointly undertake the evaluation efforts proposed here. Joint ownership of the evaluation efforts would entail close collaboration between those two agencies on the design of the evaluation, coordination of the data collection processes, and careful integration of the evaluation findings from the various evaluation efforts. Establishment of a working mechanism would be required in order to allow Teacher Corps/Washington and the Office of Planning, Budget, and Evaluation to carry out this joint ownership. The mechanism would need to include frequent meetings and a careful understanding of the desires and sensitivities of the respective agencies.

Joint ownership by Teacher Corps and OPBE would help to ensure that the evaluation was directly responsible to the evaluation needs of Teacher Corps/Washington. It would also help to ensure that considerable interaction and dialogue occur between Teacher Corps/Washington, OPBE, and the various evaluation efforts. This would lead to increased understanding of the evaluation results and a greater commitment to the evaluation efforts by the various federal agencies.

The Evaluation Task Force recommends that Teacher Corps/Washington and OPBE establish an Evaluation Studies

Reaction Panel which is representative of nationally recognized expertise in evaluation and of role groups (including community members) involved in local Teacher Corps projects. The Reaction Panel would have the following characteristics:

- It would serve as an advisory group, rather than a decision-making board or council.
- It would be broadly representative of evaluation expertise and Teacher Corps practitioner community.
- It would have a maximum size of ten members.[5]
- Appointment to the Reaction Panel would be for an indefinite, lengthy term.
- It would hold regularly scheduled meetings.
- It would receive staff assistance from Teacher Corps/Washington.

The Reaction Panel would serve a number of important roles. For Teacher Corps/Washington, the Reaction Panel would help in identifying substantive issues which should be addressed in the evaluation. It would also help Teacher Corps preserve the most useful balance between independent and internal evaluation, between qualitative and quantitative analysis, and between site-specific and cross-site analysis. The desired goal here is to synthesize the strengths of each approach. The Reaction Panel would also help in the utilization of evaluation findings by Teacher Corps/Washington. For local projects, the Reaction Panel would provide representation and a vehicle for two-way communication between the projects and other groups conducting the evaluation. It would also enhance the credibility of and cooperation with the independent evaluators which seek to work with local projects.

For OPBE, the Reaction Panel would help to clarify desired directions and methodologies for the evaluation effort. It would also help the independent evaluator, both by providing clearer insight into project processes and outcomes and by helping to clarify useful substantive frame-works which could be used in the National Process-Outcome Assessment. The Reaction Panel could review instruments and evaluation design for the study and help the independent evaluator secure

[5]RFP 78-90 calls for 20 members.

greater cooperation and collaboration from local evaluator/documentors.

The Reaction Panel need not provide all of this assistance itself. Rather, it should view itself as a facilitative body which provides process links between various cooperating institutions. For example, it could facilitate strategies for promoting cooperation and collaboration between local Teacher Corps projects and the independent evaluator or among local Teacher Corps projects themselves. This latter cooperation might take the form of regional networks of projects.

EXCERPTS FROM REQUEST FOR PROPOSALS

New projects to be funded in July 1978 will be known as Program 78. New projects to be funded about mid-1979 will be known as Program 79. At that time, there will be approximately eighty (80) Program 78 projects in their second year and forty (40) Program 79 projects in their first year. After Year One of each Program, grantees will be eligible to have their grants renewed annually for the four subsequent years. After Year Three there will be diminishing levels of Federal funds since institutionalization of the activities, using local funds, is expected to begin.

The organizational structure of the Teacher Corps is as follows:

- the National Office, headed by a Director, within the U.S. Office of Education and consisting of about 40 people in three Branches—Management, Operations, and Development;
- a system of Networks operated by Boards with contracted responsibility to identify and meet needs and provide important information links within the project community and among the several projects within the purview of a given Network;
- at the project level two groups are formed to share information, guide the projects, and provide community-based support—an elected Community Council *and* a project Policy Board. The nucleus of the Board must be the Dean of the institution of higher education, the Superintendent of the local education agency, and the Chairperson of the Community Council;

- at the local project level a collaborative arrangement between a university, a school district with typically up to four schools participating and the community served by those schools.

A typical Teacher Corps project usually includes at least one, but not more than four schools (elementary schools must be Title I eligible) within a feeder system; a team of at least four interns and one experienced teacher as a team leader; an elected community council having at least seven members; an institution of higher education; a local education agency, and other institutions and agencies such as correctional institutions, teacher organizations, professional associations, social service agencies, and students.

Orientation of Study

Basic Requirements and Assumptions

The focus of the evaluation will be on Program 78 and Program 79 Teacher Corps projects. Some 80 local project grants are to be funded for Program 78 and 40 for Program 79. Separate grants are to be made to the participating institutions of higher education, to the LEA(s), and to the correctional institution(s), if appropriate. The study design is to be longitudinal, covering the five years of Program 78 and four of the five years of Program 79. The study will emphasize the four phases of the program, i.e., Year One, Development; Years Two and Three, Operational; Year Four, Institutionalization; and Year Five, Adaptation. The activities, the processes, and the interrelationships of program participants will be the focus of the evaluation findings.

Though evaluation requires the development of a basic five-year design, the first three years will be the primary focus of the initial funding period. The second funding period will cover Years Four and Five. Grant proposals will address the entire five-year program. Therefore, some information on each year of the program will be immediately available for design purposes. More definitive plans, however, are expected to evolve as the local projects are implemented. Each local project consists of four phases covering five years.

- *Development (Year One)*—The development year will include election of the Community Council, development of specific plans for the following four years and the beginning of collaboration and involvement of the various constituencies at the local project levels.
- *Operation (Years Two and Three)*—The operation years will include training of the teachers and interns, continuing collaboration with and involvement of the various constituencies, and gearing up for and beginning to induce institutionalization of local project practices in the schools and universities of the local project. Dissemination of these practices outside the local projects may also begin in the operation stage.
- *Institutionalization (Year Four)*—The insitutionalization year will emphasize continuation of newly developed practices, procedures and materials in the organization routine of the schools and universities. This year may also include continued training for teachers and other educational personnel at the schools and efforts to demonstrate and disseminate lessons learned and products and practices developed, to schools, universities and other educational agencies outside of the project.
- *Adaptation (Year Five)*—The adaptation year will emphasize adaptation of project lessons learned and products and practices developed by schools, universities and other educational agencies not directly involved with the project. While demonstration of project successes and failures may have begun in earlier years, the outreach efforts of the local project are expected to be completed during this fifth year.

The development years (1978-79) of Program 78 and (79-80) of Program 79 will be spent by the projects in refining the goals and objectives outlined in their proposals, in hiring and organizing staff, in establishing the panels and councils specified in the Teacher Corps regulations, and in developing general management systems. The Year One evaluation activity of both Program 78 and Program 79 will be observation and examination of processes and interactions among participating groups, in the community, the schools, and the institutions of higher education, as a Teacher Corps project is being started. To accomplish this end, numerous on-site observations will be required. Reporting from this year will be formative, and should provide the basis for a good information document for Teacher Corps. It should provide, as well, an explication of the dynamics of project implementation, which may be useful for

planning and managing subsequent funding cycles.

Year Two of the evaluation will involve continuing developmental activities for Program 78 and initial work on Program 79.

Evaluation will be multi-purpose, serving the programmatic needs of Teacher Corps/Washington, local projects, and of various other levels of decision making. Teacher Corps/Washington needs information that will help it provide better service to projects, and to use better its staff resources and funds. Local projects need information to improve their operation, which will help them to become more responsive both to the needs of all their participants and to Teacher Corps/Washington. Other levels of decision-making need information that will aid in drawing inferences about future funding levels, about institutionalization, dissemination, teacher training institutes, etc. Data-gathering activities must accommodate these interests. A considerable amount of effort should be spent at the beginning of this evaluation to ascertain more clearly specific kinds of information needed. A substantive interim report from the evaluation process must be completed in time for Congressional hearings.

Both quantitative and qualitative data will be collected, analyzed and interpreted. Observational data will be collected from all 80 Program 78 projects and all 40 Program 79 projects in their first year. Trained local documentors can provide much of this information. In subsequent years extensive data will be collected from a sample of these projects, which will form the basis for the national evaluation.

The study must be performed in a manner which is responsive to practitioners. A Reaction Panel (REAP) will be formed to help formalize this effort. (See the end of this chapter.)

Focus of the Evaluation

Evaluation, both formative and summative, will be directed to Teacher Corps decision-making levels (local, regional, national), and will emphasize national level decision-making requirements. The study will also be directed towards the education profession. It will provide better understanding of the pro-

cess of implementing new ideas and practices in a field situation.

Although the study will focus on Program 78 and Program 79, it will include management information useful for future Teacher Corps funding cycles. It will focus on the different processes used to implement Teacher Corps goals, and will also focus on the unintended, as well as the intended outcomes and the impact of implementation procedures. All in all, it will provide an informed account of Teacher Corps innovation in terms of what actually occurred.

Principal Program Outcomes

Major program outcomes which are to be treated in the evaluation are:

- Improved school climate, which fosters the learning of children of low-income families;
- An improved educational personnel development system for persons who serve or who are preparing to serve in schools for children for low-income families;
- Continuation of educational improvements (including products, processes, and practices) made as a result of the project, after Federal funding ends; and
- Adoption or adaptation of those educational improvements by other educational agencies and institutions.

Primary Study Questions

The following questions are suggested. Others are likely to emerge as more knowledge of the content of the projects is collected and as specific hypotheses are developed. However, these questions are of primary interest for the evaluation.

- What are Teacher Corps projects doing to implement the key features stated in the authorizing legislation and the Rules and Regulations? What are the problems, successes and failures in the implementation of these features? Why?
- To what extent are Teacher Corps projects achieving the four outcomes stated on page 4 of this document and in the Rules and Regulations? Why?
- What is the interrelationship between the achievement of the four outcomes and the attainment of the legislative purpose?
- How do Federal laws and policy become implemented at the local

level? (How does the Federal level deal effectively with local levels?)

- What does "impact" mean in terms of legislative purpose and intended and unintended outcomes? In relationship to collaboration, institutionalization, demonstration and the implementation and achievement of the key features of the outcomes?
- What does "impact" mean in human terms, on the attitudes and behaviors of children, individual teachers, professors, community members, and other school and university personnel?
- How do Federal legislation and rules and regulations affect local projects.?
- What are local project objectives and their relationships to the national program objectives?
- What are the relationships of other Federal, State and local programs in LEA's to the Teacher Corps project?
- What are the more salient unanticipated problems, successes and failures encountered in managing and implementing specific features of the rules and regulations?
- What can be learned from Teacher Corps experience to help improve teacher training, teaching and the process of effecting change in schools and universities?
- What learning from Teacher Corps experiences can be used by other Federal agencies, in terms of implementing Federal legislation and of working collaboratively with both the local and Federal levels?

As noted from these questions, the emphasis is upon what the projects tried, and how they implemented, succeeded or failed within the Teacher Corps Program at the local, regional and national levels.

Scope of Work

The work to be performed under Teacher Corps evaluation requires the development of a five-year evaluation design. The initial funding will cover the first three years of evaluation activities. The proposed five-year design will be conducted in three phases: (1) planning, (2) operations, and (3) analysis and reporting. The period of performance under phases 1 and 2 will be 36 months from the date of award. The third phase will be optional. If exercised, it will have a period of performance of 24 months. Some activities in these phases necessarily overlap and should be so noted in the design. Project planning and development activities will take place during the first year of

the evaluation activity. Even though planning is to be the primary focus of the first 12 months of the contract, some information useful for improving program management within and between the projects and Teacher Corps/Washington is expected as a product.

Phase 1—Planning

Tasks to be performed under Phase 1 are:

1. Proposal review.

Each funded Program 78 Teacher Corps proposal must be carefully reviewed as soon as possible following the award. A review will provide information necessary for design purposes. It is expected that a wide range of strategies and objectives will be represented among the 80 funded.

2. Attendance at the initial grantees' meeting, and coordination with the Reaction Panel.

Approximately one week after the award of the contract, a meeting in Washington, D.C. is scheduled, at which key project officials representing all grantees will be in attendance. It is expected that many points about the proposals will be clarified and that some ground rules for managing the projects and for sharing information among all participants will be set. The contractor will have an opportunity to present an overview of the proposed evaluation, to meet each grantee, to establish working relationships, and to observe initial interactions among participants. The conference will last approximately five days. If the award date prohibits the contractor's attendance at this meeting, another session will be scheduled in Washington at a later time, immediately following the award.

3. Training local documentors.

Each project is expected to have one person on its staff assigned the task of site evaluation or documentation. This person will identify issues of particular interest to the project and conduct related studies. In addition to these and other local duties, this person will be able to perform valuable services for the national evaluation effort, if properly prepared. To this end, the national evaluation contractor will be required to prepare training materials and train each local documentor. Training can take place at the three regional workshops discussed below. Some training of local documentors is to take place at the national conference.

The contractor may want to spend a day or two at one of the Recruitment/Community Technical Resource Centers located in Washington, D.C., Detroit and Los Angeles, reviewing their materials on the local election processes and revising preliminary training materials.

On-site observations of such activities as the formation and deliberations of councils, community participation, IHE involvement with the project, and the interactions among participants are to be made at least three times during the year. Particular emphasis shall be devoted to the first 90 days during which time the community election will take place. Local documentors will perform much of this with spot checking by the contractor. Explicit training and reporting formats are to be developed for this task. Since Office of Management and Budget cleared materials will not be available in the first year, informal guides will be used. The contractor should be prepared to provide such materials within the first two months of the contract.

One of the important products of Year One is documentation of the project installation and implementation processes. Carefully designed observational strategies can provide useful information on the process as well as establish base line behavior data on key participants at the site level. Since local documentors are expected to play a key role in this activity, appropriate training must be provided.

Initial observational data are to be collected and reported by each grantee. To improve the value of these data to the national evaluation, observational instruments will be designed by the contractor and cleared through OMB by the second year of the contract.

4. Preparation and training in the use of different evaluation strategies.

Each project in Program 78 and Program 79 will be expected to implement a local evaluation strategy. In order to facilitate this effort the contractor will be asked to develop four different evaluation strategies from which local projects will select. Such approaches might include: (1) Discrepancy Evaluation (Provus), (2) Illuminative Evaluation (Parlett and Hamilton), (3) Responsive Evaluation (Stake), and (4) Goal-free Evaluation (Scriven). The contractor may suggest others. The contractor will develop materials or collect them from existing sources. He will prepare separate kits using four selected approaches. General presentation will be made at three regional workshops to Teacher Corps project representatives (possibly four people from each project including the local documentor), after which selections will be made. The skill of the local documentors may be a determining factor as to the kind of training provided. Once this process is completed, the contractor will provide more in-depth

training on the individual approaches in autonomous groups. It is anticipated that the workshop will cover a three-day period with separate sessions held for local documentors. It is also expected that some overall discussion will be held by the contractor to help attendees become more aware of the uses of program evaluation, thus helping to build positive attitudes toward evaluation.

Scheduling and location of the three regional workshops are to be established by the Project Officer and the contractor. However, U.S.O.E. will be responsible for securing space for the meetings and for getting attendees to the workshops.

5. Cross-site observations.

Some initial observations will take place early in the project, perhaps during the first three months. The contractor will be required to collect information from each site using an informal guide and a set of preliminary procedures, which the contractor or the local documentor may use. It is important to record some of the problems associated with implementing the proposed plans of grantees. The contractor may want to use this occasion to test some of the procedures or to develop some that may be used later in the study and to obtain information to be used for sample selection. The local documentors and the contractor may share the responsibility for conducting the initial sets of observations. Subsequent observations on site will be performed primarily by the local documentor with some spot checking by the contractor. For purposes of preparing proposals, the contractor shall assume that observations are to take place three times during the first 12 months, one set of observations by the contractor, and two by the local documentor. The frequency of this type of observation in subsequent years should be discussed in the proposals. These initial observations are designed to provide feedback to Teacher Corps/Washington and to local projects on problems of project implementation and for suggestions for program improvement. Proposals should contain examples of specific questions used in previous investigations of project implementation and installation. These questions could form the basis for the initial observations. It is anticipated that some data will be obtained from all participating groups, i.e., schools, institutions of higher education, community, and the various advisory boards.

Study Design

Formulation of Hypotheses

During the first year it is expected that few initial hypotheses will be developed and tested. Refinement of these and other

hypotheses will take place in Years Two and Three as more information about the projects is available. The hypotheses are to relate to the questions posed earlier in this document.

Sample Design

The initial observation activity will include all funded projects for Program 78 and Program 79. Subsequent data collections will involve the administration of questionnaires, personal interviews, and observations conducted in a 25 per cent sample of projects. Local documentors will provide observational data from all projects on a periodic basis, throughout the life of the contract, which will be summarized and interpreted by the contractor.

Questionnaires will be administered in all funded projects. The nine respondent groups to receive mail questionnaires are parents, policy board, interns/aides/volunteers, teachers, council, principals, local project staff, IHE faculty, including administrators, and others (state education agencies, Teacher Corps/Washington staff, etc.). At least five respondents from each group should be included in the sample. Self administration is recommended with the help of the local documentor to follow-up non-respondents.

In-depth interviews with at least one respondent from each of the nine groups will be conducted once each year for Program 78 and Program 79 in a 25 per cent sample of projects. In addition to cross checking some of the responses obtained in the questionnaire phase, the interview sample will provide further amplification on measures of awareness (Teacher Corps project), levels of participation, influence of projects, and satisfaction. Site observations will take place in the schools, the community, and the institution of higher education (IHE). A sample of local projects, schools, and classrooms within universities and schools will be selected for Level I observations. In Level I, the focus will be on school, community, and university processes as they relate to school climate variables. Level II observations will be external to the schools and universities and will include meetings of the Policy Board, of the Community Council, and of other groups associated with the project as noted in the list of nine respondent types. Level III observa-

tions will be on IHEs, including faculty, classrooms and administrators. The focus here is on obtaining more precise information about the complex process of installing and implementing a Teacher Corps project—its dynamics, interactions, and consequences on schools, IHEs, and communities during all phases of the project.

The same sample of local projects will be used for the Level I, II and III observations and for the in-depth interviews. The sample of local projects is estimated to be 20, or one-fourth of the expected projects to be funded in Program 78 and 10 from Program 79.

One design and analytic procedure to be considered in the evaluation is time series. This procedure requires a large amount of data. Consequently the sample design will include item or matrix sample techniques for a few variables. The identification of these variables will take place during the first year. It will be necessary to examine each variable to determine the frequency each must be observed in order to obtain reliable information. Further discussion of the time series approach is presented under the subsection titled "Analysis Plan."

Program 79 will begin developmental activities as Program 78 begins operational activities. Some of the preliminary findings from Program 78 may be useful in planning Program 79.

Data Collection Instruments

During the first two years of the evaluation the selection, development, field testing, and clearance of suitable instruments will be emphasized. Where possible, existing tests or scales will be considered. Three types of instruments seem likely to be used: observational, questionnaire, and interview guides. The following are examples of instruments.

1. Observation instruments.

> While educational research literature contains many important works which attempt to describe classroom environment, this study will attempt to describe the school climate. Such items as the amount of time spent on tasks, independence and persistence may be observed. Interactions between and among role groups such as teachers, children, teacher-interns, school and university teachers, parents, council members, etc., and other combinations are to be

observed—all within the total context of the community, the IHE and the LEA. Observation instruments are to be adapted or modified, where possible, from among those that are currently available.

Another important task related to development or selection of observation instruments is development of procedures for administration. As a result of this task, observational procedures and training guides will be developed.

2. Questionnaires.

A set of questionnaires is to be developed and administered at each project site. Self-administered mail questionnaires will be considered. Such items as types of educational environment, measures of attendance, volunteer hours, preservice/inservice hours, and attitudinal scales are to be included. Again, the use of existing tested items is encouraged.

3. Interviews.

Structured and unstructured interview schedules are to be developed to address such areas as: awareness of project, awareness of project goals, participation in project activities, influence of project on other institutions (e.g., schools, universities, teacher training institutions, LEAs) and satisfaction with the project. Appropriate training materials for interviews will be developed as part of this task. Some structured and unstructured questions with probes may also be considered.

Data Collection

The only data collection activity during Phase I of the evaluation will be the initial observations and the two subsequent site observations, which are designed primarily to provide information on project implementation and organizational formation and data for planning the sample of 20 projects and for providing feedback to local project management and Teacher Corps/Washington. They are to be carried out primarily by local documentors.

During the first year also, preliminary plans for collecting data in subsequent years are to be formulated, including criteria for selection of sample sites and respondents.

Analysis Plan

Much of the developmental work on the analysis plan is expected to take place under Phase II of the evaluation. However,

in designing the evaluation study, key issues of analysis must be addressed under Phase I. Year One will focus on process and formative outcomes. The analysis strategy must reflect this orientation. While some baseline data may be collected in some domains using acceptable attitudinal scales or established observational techniques, it is not likely that a great deal of this can be done in Year One.

As stated herein, the evaluation must provide information that will be useful to a wide audience. The analysis strategy must both reflect the issues and show some reasonable likelihood of providing reliable answers. The strategy will be complex enough to uncover complex relationships, yet simple enough to minimize the caveats that lead to unclear conclusions. While such a strategy may not be forthcoming, it must be attempted.

Over the past several years a fair amount of work has been done using time series to determine the impact or effect of an event. Some events in Teacher Corps may best be described through the use of time series analyses. Time series analyses have been applied in an evaluation of 1975 Corps member Training Institute (CMTI Impact Study, Fox, et al., 1976). In this study some 18 data collection points were required. Much of the work done to date has been in the univariate domain; most of the events of interest are multi-variate. Nevertheless, it is desirable to explore the use of time series analyses on a limited number of variables to be studied as part of the national evaluation. Item or matrix sampling techniques will be employed when possible along with data provided by local documentors. These sources of data may be sufficient for the time series approach.

Concurrent with the national evaluation there will be several in-depth studies conducted by highly skilled professionals examining such issues as socialization into Teacher Corps, collaboration, university climate in which Teacher Corps works, etc. This work will be available and may be useful in interpreting findings from the national evaluation.

Follow-up Procedures

Preliminary procedures for following up on selected respon-

dents are to be developed in the first 12 months of the contract and refined during the next 12 months. The data base will be designed so that new data can be linked to previous data from a respondent and so that candidates for follow-up study can be identified.

Special In-Depth Studies

This task involves two special case studies which are designed to provide further amplification of the findings produced by the basic evaluation. The special studies will be a deeper examination of issues relating to Teacher Corps efforts to fulfill its legislative mandate. The studies are to be carried out independently from the basic evaluation, by specialists in the fields of education, anthropology, ethnography, or sociology. They are to be completed within the first 18 months of the contract, so that findings can be used in interim evaluation reports. These studies may be conducted in one project or more than one. It is anticipated that the time series analysis approach will be followed where appropriate. The studies will be conducted in such a way that the findings can be integrated with those from the basic evaluation. A general description of the studies follows.

Study I—Features of Teacher Corps Implementation.

This study is to be conducted from two perspectives, *multi-cultural education* and *collaboration*.

The first perspective is to be an analysis of the cultural and multicultural milieu of a project and the implications of this milieu on project implementation and operations, as well as on the design and implementation of training strategies. This study will include an analysis of underlying assumptions held by participants in the project, of their motivations and the reasons why certain actions were taken. It will include descriptive accounts of the actions taken in developing education that is multicultural. This analysis will focus on Teacher Corps project operations in the milieu of the university, the school district and the community in which Teacher Corps projects are found.

The rationale for this aspect of the investigation is the importance in Teacher Corps projects given to education that is multicultural.

The second perspective is to be an analysis of how university, school and community collaboration is initiated and maintained and of how ownership in the practices and goals of certain Teacher Corps projects is induced. This will include a description of the

extent to which the collaboration and governance goals stated in Teacher Corps rules and regulations are implemented. It will include also an analysis of what was done, the problems that were encountered and the reasons that certain degrees of implementation were or were not achieved. This phase of the investigation may also consider the relationship between the collaborative measures taken by the Teacher Corps project and the extent of institutionalization of these measures into the university, school district, school feeder system or community. The relationship between the process used in developing the proposal and the process used in implementing the project may also be included in this investigation.

The rationale for this special in-depth investigation is recognition of the importance of collaboration in Teacher Corps change strategy and in implementation of educational innovations.

Study II—Institutional Change.

This in-depth study reflects Teacher Corps efforts to change teacher training-retraining programs at institutions of higher education and the recognized difficulty of this process. This study will provide insight into the nature of change processes that may be the most useful in university settings.

This study is to be a description and analysis of the university climate in which Teacher Corps projects work, with special emphasis on the schools of education. It will include professors' views of their practices and intentions, as well as student perspectives. (These "students" may include teachers who are gaining academic credits.) It will also include the views of university administrators such as deans and department chairpersons. This analysis and description will include descriptions of the teacher training program(s) operating in the school of education and the factors that impinge upon the design and function of these programs. It will include, as well, a description of the nature of Teacher Corps project involvement in the teacher training program(s) and the relationship of this involvement to factors that may be expected to change the program(s). This investigation will analyze specific university milieus in terms of their teacher education practices, their capacity to change and the strategies for change that are being implemented by the Teacher Corps project.

Phase II–Operations

This phase of the evaluation will start about mid-1979 and extend through the end of 1983, if the option for Years Four and Five is exercised. Several of the activities which started in Year One will continue under this phase. Site observations will

continue with emphasis placed on process and formative outcomes. Several tasks will be completed during this phase. They are:

1. Finalizing research design (1979).
2. Field testing and clearance of instruments (1978-79) for Program 78; 1980 for Program 79).
3. Training procedures for Program 79 (may adopt those developed for Program 78 with necessary modification).
4. Data collation plan (1979 for Program 78; 1980 for Program 79).
5. Analysis plan (1979 for Program 78; 1980 for Program 79).
6. Follow-up procedures (for Years Four, Five and beyond; 1979).
7. Data management procedures (1979).
8. Format for final evaluation report (1980).
9. Data collection for interim and final reports (major data collection begins in 1979.)
10. Analysis (1980-81 for Program 78; 1982-83 for Program 79).
11. Final reports (1981; 1983).

Phase III–Analysis and Reporting

The tasks to be completed during Phase III are:

1. Analysis.

The approved analysis plan developed under Phase II is to be implemented under this task. Data collected during the first three years include baseline responses, observations, interviews, and questionnaires from both Program 78 and Program 79 projects. Some longitudinal data sets are also expected to be available. Additionally, some information collected through special studies accomplished external to the national evaluation will also be available. These sources of data will form the basis for an imaginative analysis. The analysis may be as complicated as needed to explicate reliable and meaningful findings in the areas of improving school climate, program implementation, and the articulation of Teacher Corps goals and practices within and outside of the Teacher Corps domain.

Analysis for Years Four and Five will include additional longitudinal results on Program 78 and Program 79 that may address the issues of program impact more directly than the end of Year Three

analysis. The analysis under this phase will also address the issues relating to institutionalization and adaptation of the Teacher Corps program.

2. Progress Reports.

The contractor will prepare and submit a monthly progress report. This report will contain detailed discussion of tasks begun and in progress, major decisions made and their rationale, problems encountered and proposed solutions, schedule of activities for the coming month, and an expenditure report. If consultant or other planning conferences are held, summaries of these sessions are to be included in the reporting.

Quarterly status reports will be prepared and submitted to the Project Officer and shared with the Reaction Panel (see below). These status reports will be needed one week before scheduled Reaction Panel meetings. The schedule will be provided following the formalization of the panel.

Periodic feedback of information to local projects and Teacher Corps/Washington will also be required. Information relating to program operations and management will be used to improve the implementation of the projects.

The contractor will prepare and submit an annual report. *Year One*, basically, will describe the process of Program 78 start-up, planning strategies, and implementation problems. It will be possible to describe the types of programs funded and characteristics of sites. *Year Two* will provide more information from observational strategies, preliminary surveys on the optional phase of Program 78, and some comparative information on the developmental phase of Program 79. *Year Three* will provide substantial information with some measures of impact included. Since this is to be a major report of findings, considerable effort by the contractor and USOE will go into the preparation of the document. This report must be available for the 1981 Congressional hearings.

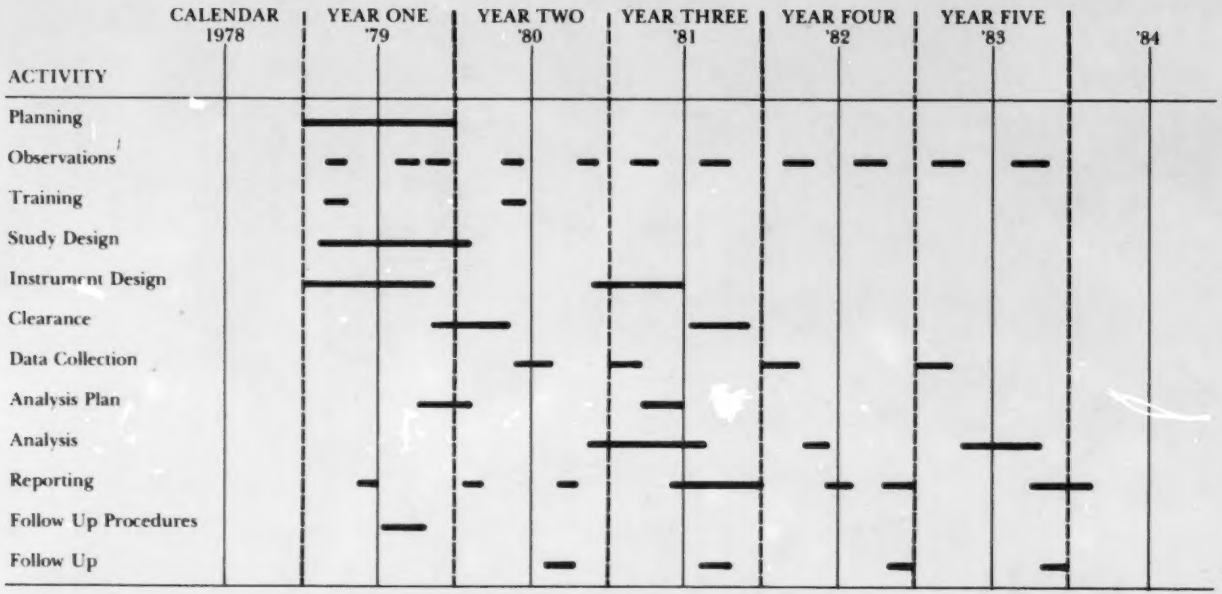
Once each year the contractor will make a formal presentation at the Teacher Corps National Conference. The presentation may take the form of a progress report.

Figure 4 displays USOE's conception of the major activities and their timing sequences over the full 5 years of the proposed evaluation.

The Reaction Panel

A 20-member Reaction Panel (REAP) will be formed by Teacher Corps, the Office of Planning, Budget and Evaluation, and the evaluation contractor. The panel will promote dialogue

Figure 4. Preliminary Schedule



among the various participants in Teacher Corps, on all activities which evaluate, document, or demonstrate Teacher Corps to a national audience. *REAP will not serve as a contract monitor nor perform clearance functions.* The panel is to be made up of representatives of Teacher Corps role groups, members of various national organizations, and individuals knowledgeable about evaluation.

Nine people will be selected by the evaluation contractor, Teacher Corps, and OPBE from a slate of 27 candidates representing Teacher Corps role groups. One candidate from each role group must be selected. The selection process will take place during the first week of the contract. The Teacher Corps role groups which are to be represented are:

- Project Directors
- Principals of Participating Schools
- IHE Faculty
- Parents
- Community Chairpersons
- Executive Secretaries of Teacher Corps Networks
- Interns/Aides/Volunteers
- Team Leaders
- LEA Research Officials

The second group of nominees will be provided by the following six associations. Each association will nominate one person to participate. The associations to be represented are:

- American Association of Colleges for Teacher Education
- American Association of School Administrators
- American Federation of Teachers
- National Education Association
- American Education Research Association
- Council of Chief State School Officers.

The third group will be selected from nine nominees, three suggested by Teacher Corps, three by OPBE, and three by the evaluation contractor who will make the final selections.

Meetings of the REAP will be chaired by the Office of Education Project Officer or her/his duly authorized representa-

tive. With respect to the evaluation contract, the functions of the Reaction Panel will be to:

- expedite the two-way flow of information among the contractor, Teacher Corps/Washington and Teacher Corps field personnel, as well as among other parties interested in Teacher Corps efforts and in lessons learned about the evaluation findings and the evaluation procedures being used;
- hear and react to major issues raised by the contractor or REAP, or suggest relevant issues that may be addressed;
- maintain awareness of the contractor to the concerns and interests of Teacher Corps and the panel members;
- provide technical reviews of instruments, means of data collection and analysis procedures prior to their adoption or use in the field.
- Recommend periodically to the Office of Education Project Officer the preparation or commissioning of papers which may be relevant to and appropriate for the role groups represented on the REAP.

There will be four meetings of the REAP during each year of the contract. Although it is suggested that these meetings occur in October, February, April, and August, either the contractor or the Office of Education Project Officer may want to have the meetings occur concurrently with the production of the quarterly reports. The four meetings, of not more than two days each, are to be held in Washington, D.C. where at least two senior members of the contract team shall be present. The emphasis of these meetings will be:

October— Recommendations for changes in Teacher Corps Rules and Regulations,
February—Reports for Congressional Hearings,
April— Recommendations for Teacher Corps National Conference
August— Review of study during past year, and future plans.

The proposal will specify how the relationships between the contractor and the panel will be maintained.

All tasks and their associated costs with the REAP meetings will be assumed by Teacher Corps and are not to be included in the contractor's cost proposal. These tasks include preparation

of agendas, invitations, meeting rooms, recording, preparation of papers, summaries and proceedings, and other such tasks which may be associated with the above functions of the panel.

Perspectives on 11th and 12th Cycle Teacher Corps Projects

ROBERT N. BUSH
and
BRUCE R. JOYCE

117

Current issues in evaluation appear to be very much alive in the plans and processes by which Teacher Corps projects evaluate themselves.

Issues in Evaluation: A Summary of the Current Dialogue

At the annual meeting of the American Educational Research Association in Toronto, in April, 1978, among the approximately 1000 titles listed in the subject index of the program, *evaluation* was the most prominent, with 35 entries. The nearest competitors were: teacher behavior 22, reading and policy analysis 19 each, and teacher education 18. Three of the five (evaluation, teacher behavior, and teacher education) are directly pertinent to Teacher Corps.

Why this concern with evaluation, which seems to be pervasive throughout the country? Many reasons have been set forth, but the current concern with "consumerism" and its corollary, a call for accountability in all walks of life, head the list.

As befits its prominence, Teacher Corps has become much-evaluated, both nationally and locally. Having had eight substantial evaluations of the national program during its twelve years of existence, it may be one of the most evaluated of all federal education programs, and its concern with project-level evaluation is as great as it has been at the program level.

It is clear, as Teacher Corps moves into its new five-year project cycle beginning in 1978, that evaluation will continue to play a key role, though a somewhat different one than in the past. The length of the cycle makes it imperative that local and national evaluation be coordinated to provide for continuous fine-tuning of projects and policies. In September 1977, in recognition of this fact, Teacher Corps/Washington appointed

three Task Forces, one directly on evaluation, to assess the problem and to make recommendations for the future.[1]

The report on evaluation emphasizes the basic importance of coordination of evaluation at the national and local levels. Some members of the Task Force believed that local evaluation should receive the dominant emphasis and that considerable emphasis be directed mainly to formative purposes, namely, the continued improvement of local projects. It was because of the recognition of the key importance of integrating local and national evaluation efforts that Teacher Corps commissioned this study of evaluation in the 11th and 12th cycle projects. The aim is to search that experience, to see what it yields for the series of programs beginning in 1978, and to provide some immediate feedback to the 11th and 12th cycle projects for their use in strengthening their current evaluation efforts.

The method as been to solicit from each project documents describing their evaluation efforts, seek response to a brief questionnaire, and to conduct telephone and on-site interviews and discussions in a small sample of sites. These data were gathered in February and March of 1978.

The response was good (four-fifths of the projects sent materials) and the information enlightening. We obtained usable and dependable results, Busy project directors and colleagues responded at a time when there were many other unusually pressing concerns. For this we are grateful. We hope that the results of the survey will be useful to them. Analysis indicates that the centers responding are representative geographically of urban and rural areas and of 11th and 12th cycles.

The resulting data are meaningful not only in the context of Teacher Corps but in terms of issues currently being raised about evaluation in general.

A Brief Historical Note

Evaluation today is an extremely broad, wide-ranging field of activity. It was not always thus. As a distinct field in education it

[1]Teacher Corps Task Forces, *Document I—Reports on Demonstration, Program Evaluation and Technical Resources*, and *Document II—Materials on Technical Resources and Program Evaluation*, Developmental Training Activities, Teacher Corps, Center For Urban Education, University of Nebraska at Omaha, 1977.

began to emerge in the 1930s and thus is only about fifty years old. Wrightstone stated in the 1950 edition of the *Encyclopedia of Educational Research* that

> Evaluation is a relatively new technical term, introduced to designate a more comprehensive concept of measurement than is implied in conventional tests and examinations. Monroe has distinguished between measurement and evaluation by indicating that the emphasis in measurement is upon single aspects of subject-matter achievement or specific skills and abilities, but that the emphasis in evaluation is upon broad personality changes and major objectives of an educational program. These include not only subject-matter achievement but also attitudes, interests, ideals, ways of thinking, work habits, and personal and social adaptability. From the point of view of its functions evaluation involves the identification and formulation of a comprehensive range of the major objectives of the curriculum, their definition in terms of pupil behavior, and the selection or construction of valid, reliable, and practical instruments for appraising the specified phases of pupil behavior. The instruments of appraisal include achievement, attitude, personality, and character tests,rating scales, questionnaires, judgment scales of products, interviews, controlled-observation techniques, sociometric techniques, anecdotal records, stenographic reports, and sound recordings. In addition, evaluation includes integrating and interpreting the various indexes of behavior changes into an inclusive portrait of an individual or of an educational situation.
>
> Curriculum making and evaluation are integral and interacting parts of the educational enterprise. Truly comprehensive evaluation provides evidence of the degree to which important curricular purposes are being realized. This evidence may lead to new curricular policies which may, in turn, provide new or changed objectives to be evaluated by new methods or techniques. Hence, evaluation requires the cooperation of both school personnel and test technicians.[2]

The article in that issue of the *Encyclopedia* was only three pages in length (in contrast to seventeen pages on achievement tests) and the earliest reference cited was 1932, interestingly by Ralph W. Tyler, considered by many to be the father of the evaluation movement in American education, who, although several times retired, is still vigorously at work in the field.

Evaluation represented an effort to break away from the con-

---

[2]Wrightstone, J. Wayne, "Evaluation" in Monroe, Walter S., Editor, *Encyclopedia of Educational Research*, New York: The Macmillan Company, 1950.

fining testing and measurement fields. It received major prominence as part of the Progressive Education Reform of the 1930s, in the evaluation of the Eight-Year Study, which was headed by Tyler. This study contributed significantly to that break from the confines of measurement by broadening the objectives to be given consideration and by construction of imaginative new and unorthodox means for measuring them. How far this early broadening impetus has taken us was illustrated by Webster and Stufflebeam in the 1978 AERA State of the Art Lecture in which they outlined thirteen different types of evaluation, ranging from those of political orientation, across those built around specific objectives and questions to be answered, to those that lie wholly in the "value" area. How far we have come may be seen in the value category where the position of Scriven, Stake and Eisner suggest that objectives—the heart and soul of the movement as it began—can now be eliminated and replaced by a "goal free" process that rests mainly upon qualitative matters after the fashion of connoisseurship. In the manner of art and literary criticism, their aim is to determine the fundamental worth or value of something. Where can and will evaluation go from here? Or has it run its full course?

It is little wonder then that, when we field-tested the small questionnaire for this study, informants said "You had better define what *you* mean by evaluation, as a prelude to your questions." In heeding their advice, we took the broadest possible definition as "anything that you, the respondent, might think could or should be labeled "evaluation," (not being willing to exclude anything in our initial sweep). But it would seem desirable and necessary, if evaluation efforts are to be of maximum use to Teacher Corps at this point in its history, to provide a definition of what evaluation might be and how it might best help to move Teacher Corps forward toward its mission. Borrowing from newspaper reporters, we shall attempt to sharpen our view of evaluation today by posing a set of questions: WHY? WHO? WHAT? HOW? WHEN? WHERE? and COST?

We begin with a description of the issues involved in each of the major questions. This description will be followed by presentation of the data from the survey as they illuminate these

questions—where we now are. In a final section we will deal with possible and desirable directions for evaluation in Teacher Corps.

Key Issues, Questions, and Problems That Need Consideration

Why?

This question breaks in two: (1) *Why,* in the sense, "Wouldn't it be better to use our energies otherwise, such as operating the program in the best possible manner?" and (2) *Why,* in another sense, related to the purposes of evaluation. Does evaluation emphasize the *summative* (to determine how well the enterprise has been doing on some absolute or relative scale) or the *formative* (the view which holds that evaluation is for the improvement of the operation)?

To answer the first *why,* there seems to be little choice. Strong social forces of accountability, reflected in legislative mandates that tie funding to evaluation, seal our fate. We have no alternative. The advice to relax and enjoy it might well be heeded. But on the second issue, to emphasize the formative or the summative, there is much more room to move.

The pressure from the policy and legislative levels tends to be for summative efforts. Is this program better than that one? Should it be continued or stopped? At the operational levels, where professionals and citizens work in the projects, the emphasis tends to be in quite a different place. How is this working? How can it be made to work better? Can this process be made smoother? Would something else work better? Timely, practical, and immediate feedback is valued.

The issue is what emphasis is likely to be most productive, to fit best, at different levels and under different circumstances and at different times. Can a given project attempt evaluation for both summative and formative purposes without serious negative consequences? Should it try? Or would it be better to concentrate effort on one or the other? These are some of the questions that projects must consider if evaluation efforts are to be as effective as possible.

Who?

The issue of who should evaluate may be sharpened by asking

four questions:

Should it be an insider or an outsider?
Should it be the producer or the consumer?
Should it be an "expert" or not?
Should it be all or some combination of these?

On the first question, strong but sharply divided opinions obtain. The policy makers not only favor evaluation for summative purposes but also tend to prefer outside, independent evaluation, selected and paid by and reporting to someone not connected with the project. How else, they ask, can objectivity, honesty and courage be assured?

The view usually taken by those closer to project operations prefers evaluation to be an integral part of the operation, conducted by those who work on the project—the insiders—as a regular part of their duties. Project personnel tend to the view that outside evaluators never stay long enough to become acquainted with the details of what is going on. Hence, outside assessments are likely to be superificial at best and even quite wrong. Insider-collected information is more likely to be used than is outsider-collected information.

The second question, of producer or consumer,[3] has some of the same ideas just enumerated on the insider/outsider issue, but with a slightly different cast. The consumer view grants, may even insist, that the project itself engage in continuous evaluation to improve its operations, but it insists that the *bottom line* must be those at whom the project or program is aimed, e.g., pupils in an elementary school reading program, or teachers prepared in teacher training institutions. Unless the program can show growth and satisfaction on the part of those who are to be trained and educated, the essential ingredient in evaluation is missing. The principal, perhaps even the only way, is to ask the consumer for his view. Not so, argue those who insist that the producers are the more effective evaluators. They are more expert than the consumers, they know where the real problems lie, how the enterprise works, and can tell

[3]Producer is here defined as those who operate or conduct the training (the project personnel), and consumer refers to those who are to be trained by the program.

better, if they take time and effort, how effective the operation is and how it might be improved. Resources spent by the producers, the operators, in evaluating their own efforts will bring greater benefit than those spent on consumer satisfation studies, although these may have some place in a total evaluation scheme.

On the third question, concerning the use of "expert" evaluators, seldom does anyone explicitly take the view that no expertness is needed in evaluation. But there is mistrust of those who have high technical competence and sophistication in evaluation and become so preoccupied with the minutiae of their speciality and the elegance of their designs that they miss the central thrust and substance of what the program is trying to accomplish. They fail, in the proverbial sense, to see the forest for the trees. Even when expertness is taken for granted, argument may arise concerning what *kind* of expert should be involved. Should it be a test-and-measurement-psychologist with specialization and a high degree of objectivity in his approach, or an anthropologist who can make an in-depth ethnographic case study that is highly qualitative in nature? The issue of "Who?" is so crucial that a paraphrase of one major report by a highly-placed government educational advisory committee concludes, not altogether without tongue in cheek, "Tell us the results you want beforehand and we can select the evaluator who will produce what is wanted."

Much hangs on the decision of who is to evaluate, both with regard to the reliability and validity of the results, as well as to their credibility.

What?

The terrain of what is to be evaluated is dominated by two dimensions of focus: process and product. On their relative attractiveness opinions diverge and fashions change. It is a variation of the perennial means/ends question. There seems to be a tendency for evaluators to stray from the product toward process. Process is seductive. It is filled with human events crying out to be looked at. Many fascinating byways invite scrutiny. The curriculum, the methods, the materials, the interpersonal relations are readily available to be studied. They are often

quite tangible. You can get your hands on them. And so evaluation schemes grow and grow, sometimes until the original target becomes obscured or lost. It is further argued by those enamoured of process evaluation that the essence of the enterprise may be sensed only or at least mainly through the ways in which it is operated—not your stated goals, but what you do. "Actions speak louder than words." The advocates of "goal free" evaluation appear to epitomize the view that the ends bury themselves in the means, can be inferred from the means and seen more clearly in the process. But inevitably, someone sooner or later—the realist, the hardnosed pragmatist, comes up with the question, "Yes, but what about the product?" It is extremely embarrassing to have nothing to say on the subject.

The relative attention to process and product is an important question that warrants much pondering by those who design and carry out educational evaluations.

How?

Of the questions posed thus far, this one is the most complicated, complex, and troublesome. And in many ways it encompasses much that is contained in the others. Some of the more persistent issues are:

To what extent should evaluation be relatively objective or subjective?
Shall it be goal-oriented or goal-free?
Shall it be a separate enterprise or so firmly embedded as an integral part of the enterprise as to be virtually indistinguishable from operation?

That the evaluation community is far from agreed on some of these issues is reflected in the fact that one of the two invited addresses of Division D, Measurement and Research Methodology, of AERA at its 1978 meeting was a debate between Ellis B. Page and Robert E. Stake on "Should Educational Evaluation be More Objective or More Subjective?" Is it symptomatic of a swing away from objectives and goals as the anchor for evaluation that Michael Scriven, the chief spokesman and architect for goal-free evaluation, was principal

speaker and discussant at five different sessions of that same meeting and that he is president-elect of the Association?

On the separate/integral issue, the advocates of summative, objective, and external evaluation are inclined to argue that the evaluation should be quite separate and removed from everyday operations. Evaluation should have its own budget, personnel, and space, and be free to move in and out as required, with ample time free from distraction to converse, develop instruments, and ponder the meaning of the data gathered. The opposite view is that evaluation should be so woven into the warp and woof of the project that it is indistinguishable from other aspects. More of this view will be discussed later under the question, *When?*

Is the current ferment in the field of evaluation the foreshadow of a new field, yet unnamed, about to spin off as evaluation did over 50 years ago when it attempted to free itself from the then considered too confining field of tests and measurements?

Many methodological questions have not been settled. New ones arise as fast as old ones are settled. Procedures have not yet been developed for measuring some of the more complex and intangible outcomes. Nonetheless, by comparison with even a few years ago, the current situation is a cornucopia of opportunity. A rich array of well-developed ideas, methodologies and instruments (as well as many half-formed ones) are available to those who wish to evaluate educational programs. The challenge is to determine clearly and realistically what is wanted and then to choose accordingly.

When?

When should evaluation be done? At what point in the life of the project or program? We assume that all of the projects being here considered are to be evaluated. It is theoretically possible that under some circumstances a project or program would be so brief, so clear cut in its plans, and its outcomes so evident, that no expert evaluation program would be required. However, no example comes to mind.

The possibilities for the time of evaluation include:

Evaluation toward the end or in a "pre-post" design only.
Evaluation at periodic intervals.
Evaluation continuously from the beginning.

One of the most widespread laments of advocates of evaluation is, "Too little, too late." "Why," they plaintively ask, "is evaluation always an afterthought, something tacked on at the last minute and at the end?" Teacher Corps aside for the moment, it is very often true that evaluation comes toward the end or after a project is completed. This is modal practice. However, there is a pressing logic for the "evaluation at the end" position. How can evaluation be done until after that which is to be evaluated has taken place? What would there be to evaluate in the beginning? If it is so wrong, why does almost everyone do it this way?

The periodic evaluation position argues, "No, don't begin on day one, but start soon thereafter to take soundings and continue to do so at regualr intervals to see how things are going." The logic is: "Don't postpone until the end, or you may find yourself so far off course that correction becomes impossible."

"Do it all of the time," the enthusiast advises. "Begin on day one and don't lay down the cudgels until all is finished." The reasoning is that a responsible undertaking should be continuously concerned with how well the project is going and everything possible should be done along the way to see that improvements, adjustments, and fine tuning take place. The view is that formative, process evaluations become indistinguishable from operating an effective program.

Where?

The alternatives of where evaluation should or does take place are straightforward:

Off site
On site
Both

Evaluations of all three kinds abound. Survey research based on mailed questionnaires has a widespread appeal to researchers who want large numbers and random samples for purposes

of generalization. In this case, the evaluator may never visit the sites to be evaluated, although they may visit non-project sites to get a "feel" for what they are later to evaluate and for instrument development. In addition, evaluators of this type may rely heavily on documents supplied to them from the sites. They claim greater objectivity when they examine documents rather then become involved with project personnel, who may co-opt them or otherwise entrap them with the subjectivity of feelings that interpersonal relationships engender. This type of evaluation lends itself to the summative end of the spectrum, but is often used for formative purpose as well.

The contrasting view holds that realistic, valid evaluation requres intimate and continuous connections with the site—a live-in arrangement, as it were, that has the flavor of anthropological field work. Such thorough and intimate relationships require time and, hence, sacrifice number and sampling facets, but this is more than outweighed, it is argued, by the rich detail of project events that becomes available.

The best of both possible worlds is envisaged by those who argue for both on and off-site work. One approach uses broad surveys as step one, followed by a selection of sites or situations that are thus revealed to be potentially productive for intensive case study in the field.

Where any particular evaluation might fall in this three-fold categorization needs careful attention. It is likely to be a function of the purposes to be achieved, the resources available, the biases and competence of the evaluators, and the conditions present at the site.

Cost?

How much should an evaluation cost, and who should pay? A ready answer springs to mind: as little as possible and whoever is willing! But that scarcely does justice to the problem. The question of "How much?" really means coming to grips with what should be the magnitude of effort devoted to evaluation in relationship to the total effort dedicated to the project. One frequent complaint from project personnel when they discuss evaluation is that they look upon time spent in evaluation as being at the expense of making the program work. An undue

amount of attention to evaluation detracts from doing the central and essential work of the project. If a program or project were broken into the following parts, would there be some optimum for each?

Planning and designing.
Developing of materials, instruments, treatments.
Operating the program.
Evaluating the results.
Using the results of evaluation for dissemination and institutionalization and for correction and improvement of the program.

It is our impression that not much thought is typically given to the problems of time and effort allocation in these terms. What happens, "just grows." The notion we have of the real world of project and program life is this: you work like crazy and almost kill yourself getting a proposal ready by the deadline and cranking up the process—that is category 1. Then most of the effort thereafter, if the application is successful, goes to #3—to make it work, with some modest attention to #2 beyond what was done during #1 in developing the proposal, and as little as necessary on #4 and #5.

Is this practice the best way to achieve maximum results? Or might some redistribution of effort be considered? Would an increase in time and effort and dollars in categories 2, 4, and 5, which now receive only minor allocations, increase overall productivity and make #3 much easier to perform?

On the matter of who should pay for evaluation, the often-heard remark at the project level is, "If *they* want it done, let *them* pay for it." The unspoken sentiment is, "Not any of my hard-earned money which should be spent for program operation should be taken away for evaluation that someone else wants or insists on."

Who pays for evaluation may well be a critical factor in how it is done and whether the results are used. As the Task Force on Evaluation, referred to earlier, came increasingly to be aware, *ownership* of the evaluation is a matter of central importance. A persuasive case can be made for the proposition that when ownership lies outside of and is not felt by project or program

personnel, the chances of evaluation having any effect on projects is likely to be slight.

Two of the biggest problems facing evaluation are the charges of lack of validity and reliability, and even when these conditions are met, the failure to use the results. Beyond lack of validity and reliability, what are some of the other reasons for failure to use results? Timeliness (the results came in too late!) is one reason often given. Another, which we suspect to be a strong force, is the lack of feeling of ownership. "This is not of our doing and making. It is a lay on." If this be true, and central policy makers want more and better evaluations and their use for program improvement, a way lies open to them of making "add on" money available for those who wish to make evaluation an integral part of their projects. This would require some joint responsibility for designing such evaluations and liberal provision of technical assistance to projects to enable them to carry out their plans. But we are getting ahead of ourselves, encroaching on the last section of this report which outlines possible directions.

Summary

Thus, evaluation in Teacher Corps operates in a general context in which the nature of evaluation and its purposes are widely debated both in Teacher Corps and the wider education community.

Concepts of evaluation have broadened from issues of measurement to comprehensive views which include the documentation of many aspects of program elements and the examination of a wide range of outcomes.

Formative and summative evalution compete for emphasis.

The roles of insiders and outsiders, experts and participants, are debated.

Process and product compete in both the formative and summative domains.

The virtues of objective and subjective sources are debated along with goal-oriented and goal-free approaches.

The relative merits of periodic, continuous, and post-project plans are argued.

On-site and off-site methods are compared for cost and fidelity.

Sponsorship and ownership entail substantive as well as political questions.

The Data From This Survey: Where We Are

The following data are mainly from questionnaires and written documents, supplemented by limited interviews and discussions.

The "real world" is at best dimly revealed in the written word, and few documents fully reveal the flow of events that actually happened. Nonetheless, we believe that some relatively clear outlines of the reality of Teacher Corps evaluation have appeared and that a number of meaningful generalizations can be drawn from the evidence.

In the context of these general issues about evaluation in educational programs, we now turn to a report of the results of our inquiry concerning how evaluation is being done in 11th and 12th cycle Teacher Corps projects—to an analysis of this particular "real world" and how those with responsibility view the problems they confront. After this analysis we speculate about the meaning of these data for the future.

Quality of Response

Altogether, responses were received from 100 projects. Two thirds of these returned the questionnaires. Most of the remaining third had sent materials and/or letters before the questionnaires were sent, making it redundant to their prior replies.

It appeared that the quality of response to the questionnaire items was acceptable. About 90 per cent of the respondents answered all of the items errorlessly. In some cases respondents indicated that questions were not phrased to fit their situation exactly and wrote in changes and otherwise brought the questionnaire into conformity with the configurations of their local situations.

The interviews conducted with project personnel appeared to validate the questionnaires in the sense that differences in responses were reflected in the descriptions provided by the interviewees.

Questionnaires and Documents: How Congruent?

The congruence of the questionnaires and documents was

encouraging. Sixteen projects were selected which had responded to the questionnaire and had sent in relatively complete materials describing their evaluation plans and follow-through.

The costs, personnel used in the evaluation, and the focus of the evaluation plans were examined to determine whether the responses to the questionnaire were similar to those reported in the other materials.

In all cases, the descriptive materials were congruent with the responses to the questionnaire.

The General Picture of Evaluation

Questionnaire Results

The official documents establishing the grants under which the Teacher Corps projects were operated are quite informative (can take in considerable information) about the evaluation plans. All of the 12th Cycle and 90 per cent of the 11th Cycle projects described evaluation activities in their proposals. Quarterly reports describe the evaluation activities which have been undertaken in three-fourths of all projects. (83 per cent of the 11th Cycle projects.)

About half the projects use outside evaluators, with somewhat more (61 per cent) of the 12th Cycle projects doing so than the 11th Cycle (43 per cent). About four-fifths of the projects use formal means of evaluation internal to the project in the form of questionnaires, interviews, and routine data collection activities. Also, about four-fifths of the projects report that results obtained from the evaluation effort have been used in some way, generally in fine tuning the projects, in modifying the activities on the basis of discussion of the results. A few more of the 11th Cycle projects (89 per cent) report this than do the 12th Cycle projects (79 per cent), reflecting the longer period of time that they have had to receive and utilize the results.

Table 1 compares the 11th and 12th Cycle projects with respect to the descriptions of evaluation, outside contracts and internal procedures, and the presence of results that were utilized.

TABLE 1

Comparison of Eleventh and Twelfth Cycle Projects in Selected Variables About the Nature of Evaluation

(as reported in the questionnaire)

(n-70)

Is evaluation described in the proposal?	Eleventh Cycle	90%	
	Twelfth Cycle	100%	
	All Projects	95%	
		Yes	No
Do quarterly reports describe evaluation activities undertaken?	Eleventh Cycle	83%	17%
	Twelfth Cycle	62%	38%
	All Projects	74%	26%
		Yes	No
Is evaluation contracted to individuals or agencies "outside" the project?	Eleventh Cycle	43%	57%
	Twelfth Cycle	61%	39%
	All Projects	51%	49%
		Yes	No
Does the project employ "internal" formal evaluation procedures?	Eleventh Cycle	80%	20%
	Twelfth Cycle	88%	12%
	All Projects	83%	17%

Dollar Cost of Evaluation

The respondents stated that it is difficult to estimate the proportions of time being put into the evaluation programs partly because evaluation responsibilities tend to be divided among several people, and the amount of time devoted during any given period appears to fluctuate widely. Thus, where there is a contract which provides, say, for 25 per cent of an outside evaluator, there may be a variety of fractions of other individuals devoted to evaluation so that the actual full time equivalent in evaluation is close to one or two persons.

The amounts reported to be spent on evaluation fluctuated quite widely. The means for both 11th and 12th Cycle projects were in the neighborhood of $12,000 a year, but these means were pulled up by a relatively few projects which spend from $35,000 to $45,000 a year directly on evaluation. Most of the projects appear to be spending more in the neighborhood of $6,000 or $7,000 per annum. In other words, the average project makes a relatively modest investment in the evaluation process while a few use considerable resources. Also, as in the case of personnel, total amounts actually expended on evaluation activities are extremely hard to determine because of the extent of informal activity and in-kind services provided by many staff members.

A rough estimate would be that 5 per cent to 10 per cent of the project energies have been put into evaluation. Given the fact that about 8 out of 10 projects have been able to use some aspects of evaluation to modify their programs and policies, this percentage would seem to be fully warranted.

The Focus of the Evaluation of Activities

Table 2 compares the focus of evaluation in the 11th and 12th Cycle projects. Generally speaking the focuses were very similar with some differences appearing that are apparently a result of the stages which the projects have reached (particularly that the 11th Cycle projects are in their second year and the 12th in their first.)

Over nine-tenths of the projects reported that they devoted "quite a bit" or "extensive" energy to the documentation of planning and governance processes although the emphasis was somewhat stronger in the 12th Cycle where 36 per cent of the projects reported extensive focus on process. Similarly, more than nine-tenths of the projects put "quite a bit" or "extensive" energy into the documentation of training and other project events. This was in fact the area receiving the greatest amount of attention by all projects in both cycles. Ninety per cent of the 11th Cycle projects reported "quite a bit" or "extensive" emphasis on the documentation of training and other events and 96 per cent of the 12th Cycle projects did so.

Documentation of teacher and intern growth and classroom

TABLE 2
Focus on Evaluation Activities
by 11th and 12th Cycle Projects
(in percent of respondents)

Planning and Governance Process	Not at all	Some	Quite a bit	Extensively
Eleventh Cycle	10%	29%	42%	19%
Twelfth Cycle	10%	27%	27%	36%
All Projects	9%	28%	37%	26%
Documenting Training and other Project Events				
Eleventh Cycle		10%	42%	48%
Twelfth Cycle		4%	61%	35%
All Projects		7%	50%	43%
Documenting Teacher and Intern Growth in Classroom Skills				
Eleventh Cycle		26%	42%	32%
Twelfth Cycle	9%	26%	43%	22%
All Projects	4%	26%	42%	28%
Documenting Attitudes of Clients				
Eleventh Cycle		33%	50%	17%
Twelfth Cycle	9%	23%	64%	4%
All Projects	4%	29%	56%	11%
Documenting Teacher Corps Staff Skills				
Eleventh Cycle	13%	52%	29%	6%
Twelfth Cycle	22%	48%	22%	8%
All Projects	17%	50%	26%	7%

TABLE 2 (Continued)
Focus on Evaluation Activities
by 11th and 12th Cycle Projects
(in percent of respondents)

Pupil Learning	Not at all	Some	Quite a bit	Extensively
Eleventh Cycle	25%	40%	29%	6%
Twelfth Cycle	32%	36%	27%	5%
All Projects	28%	42%	28%	2%
Organizational Change				
Eleventh Cycle		43%	43%	14%
Twelfth Cycle	8%	33%	46%	13%
All Projects	4%	39%	44%	13%
Institutionalization				
Eleventh Cycle	3%	32%	45%	20%
Twelfth Cycle	4%	30%	62%	4%
All Projects	4%	31%	52%	13%
Evaluation of Materials				
Eleventh Cycle	16%	29%	42%	13%
Twelfth Cycle	23%	45%	23%	9%
All Projects	19%	36%	34%	11%

Has the Project had results from the evaluation efforts	Yes	No
Eleventh Cycle	89%	11%
Twelfth Cycle	79%	21%
All Projects	84%	16%

skills was also reported to receive "quite a bit" or "extensive" attention with about two-thirds of the 11th and three-fourths of the 12th Cycle projects providing "quite a bit" or "extensive" attention to that area.

Finding out how the clients feel about the program was also a strong focus of most of the projects with over two-thirds giving it quite a bit of extensive attention.

Documenting Teacher Corps staff skills received some attention but less than the above categories, although about a third of the projects indicate that it received "quite a bit" or "extensive" attention.

Pupil learning also received moderate attention by about a third of the projects, and none by about a third.

A few of the projects planned for comparisons between pupil achievement and the project schools on pupil achievement and others, but the norm was not to make a considerable investment in the area of pupil learning.

Organizational change, however, received moderately high emphasis with nearly 60 per cent of the projects giving it considerable attention. Institutionalization received considerable attention also. Both of these areas have been recent emphases of Teacher Corps projects and are reflected in the focus of evaluation.

The evaluation of instructional materials receives only moderate attention with nearly a fifth of the projects giving it none at all and about a third indicating that only "some" attention was given in this area.

On the whole, the two cycles were quite similar with respect to the focus of evaluation with the chief differences being in areas requiring time for development. In those areas the 11th Cycle projects reported greater development. However, more of the 12th Cycle projects used outside evaluators who are contracted for the purpose. In addition, the older projects give somewhat more attention to pupil learning (but not a great deal), and somewhat more to institutionalization.

Use of Data for Decision Making.

Projects reported that the primary internal use of evaluation was for getting a picture of how the project was doing and for

adjusting the program (formative evaluation).

> Internal usage was for project ongoing feedback and adjustment.
>
> . . .
>
> We use a constant feed-in from clientele analyzed during the three to five day staff meeting. The result is refinement of the developing teacher education curricula and also identifies role education needs not being well met.

Another project reports a variety of uses:

> Our activity logs have been helpful in reflecting overall trends in our activities. For instance, we have learned that almost three-quarters of our Teacher Corps activities have been on site in accordance with our proposal. Also, the logs have reflected a change from planning to implementation type activities.
>
> Our clinical team evaluations have been used to feed back to teachers on the nature of their interactions.
>
> Assessment has also been made of the extent to which teachers have gained competencies in content and methods in reading and mathematics.

Some other projects have identified particular ways in which they have changed:

> Some staff reorganization with a new role added was seen as essential.
>
> It is helpful that we expect the change and use the informal results to make decisions about changes.

The informal theme continues throughout. Apparently one of the results of the emphasis on evaluation, albeit is relatively minor, has been to sensitize people to the utilization of data for change.

The relationship between the utilization of evaluation and the type of structure developed in the project was underlined repeatedly.

> We have formed a needs assessment and evaluation leadership team consisting of eight teachers and monitor the inservice program. Based on periodic data collection and analysis we have modified our project's operation and activities.

Another comment:

> Informal results which have led to project modification and delivery services.

A few of the projects have made formative use of student achievement data.

> We have achievement scores on each student in the project for the past three years. Each teacher gets a readout on her class which also gives self-concept scores, attitude toward school, etc.

That particular program utilizes a computerized system providing information about achievement and attitudes directly to the students on a periodic basis.

In many cases the projects have used the results for internal purposes. One project, for example, described explicitly using evaluation results with the Dean in order to persuade him to continue the project, and another described using them with the local education agency. Altogether, about a dozen projects reported explicit use in their political milieu aside from the information that was used for projects, project management purposes, or reported to Teacher Corps in general.

Although quite a number of projects, as can be seen above, described relatively comprehensive objectives for the evaluation component of the projects, implementation, obviously, was quite another matter. In several places a complex, almost mechanistic overall design was followed by implementation of a different sort. For example, one of the projects which included in its proposal a relatively complex plan for systematic data collection and analysis presented the following advertisement for the evaluation consultant:

> The evaluation consultant will operate as a modified participant observer during the project's operation. This will entail observing in-service activities in the school district, interviewing a full range of project participants, observing staff activities, and writing four reports to be part of the total evaluation demonstration materials for the project.

Thus, the major investment of this project (which spent about $9,000 altogether on evaluation) was on a more descrip-

tive type of evaluation than the original plans would have indicated.

The Study of Side Effects. A few projects built in attempts to measure side effects of their activities.

There were, however, some outstanding examples among them. In one case a project generated a series of case studies of students to determine whether their lives had generally been affected by events in the school in which the Teacher Corps was centered. These case studies were made of a very few students selected at random from the four grades of the high school in which the project was centered and provided an excellent indication of whether the students were becoming more purposeful, better adjusted, and how they were reacting to activities which had been generated by the Teacher Corps project itself.

Frustrations, Difficulties, and Problems. The projects were asked directly about problems that were encountered. Many of the respondents were extremely sophisticated. "No major problems except to mention that attitude changes are, as you know, extremely difficult to measure in a meaningful way . . . This aspect of evaluation seemed to emerge quite consistently." And, in a different vein: "Trying to tie in a microcomputer system for evaluation was very difficult. We understand the amount of time needed. We are still not to the operational stage."

This last comment indicates the relative difficulties encountered at programs which took seriously the attempt to build in "systems" evaluation procedures that would provide continuous feedback.

It would seem unnecessary for each project to have to develop all of the software necessary to do this kind of thing and it is not surprising that comments like the above would emerge. It is worth noting that that is an 11th Cycle Project which is in its second year of operation and has invested nearly $35,000 each year in an attempt to build a sophisticated and contemporary system.

Another problem that was reported frequently relates to the contemporary need for data on the part of managers and the burden it places on respondents.

"Particularly at the LEA . . . but also at the IHE . . . community level

participants often feel 'over-surveyed' . . . caught up in the *paper chase.*"

The same project also mentioned that it is difficult to find consultants who had a "visceral" feel for the situation. The outsiders in many cases tend to remain the outsiders and the kind of data they provide may not always be useful to persons who are caught in the problems of developing meaningful evaluation programs.

Consequently the respondents reported an attempt to do an honest job, but the problem of doing it within the time and resources necessary—"total time needed to evaluate adequately *all* aspects of the project . . . we are forced to be very selective in evaluation." Throughout there was the concern with the development of useful formative evaluation as against the more familiar summative designs.

The resistance of personnel is a case in point. Projects reported that many personnel do not like to undertake the mechanical records of performance that are necessary in order to have a complete base of data about the activities on which to base an ongoing evaluation. In the words of one director, "Documentation requires a 'Boswell' on the staff, someone who can interact easily with staff members and clients."

The same project reported that it is easier to find someone who is interested in management by objectives than someone who is interested in action research and informative evaluation. "The nature of Teacher Corps project work requires flexibility and serendipity."

It was plain that quite a number of the projects had taken very seriously the problem of building systematic evaluation of both the formative and summative sort, but that the complexity was a continuous problem. The resources alloted to evaluation were simply not sufficient for so comprehensive an approach. As one project director said, one whose project has an extensive and elaborate program, "It took so long to get to this point."

Another director said, "Project teams should be trained *very* early on."

Many projects seem also to be seriously concerned with the problem of relaying evaluation information to multiple audiences and found the need to "describe evaluation procedures to

the world" taxing and time consuming. "The time to carry out all evaluations and research activities along with other project responsibilities," was cited especially in those programs which had developed complex evaluation plans. Some felt that the national office could help in this: "I think that new directors should have a much more intensive instruction in the national office on what is and is not evaluation so that we can be more productive of what we develop."

The projects also found that some of their clients needed extensive training. They cited—"lack of understanding of the evaluation process and the necessity for it among LEA teachers." This problem extends to time needed by project personnel to complete evaluation instruments.

One director suggested that Teacher Corps "place in each project a microprocessor with enough storage and have it programmed for information that Washington, D.C. needs." Teacher Corps/Washington could then call each project's microprocessor and gather the information on all projects within a few days.

Another project stated: "We have had to develop a model from scratch." translating complex models like Stufflebeam's into language and procedures that all role groups in the project could understand.

The problems of learning to use computers were cited frequently along with optimism about what can be done with the smaller contemporary computers. However, with all the frustrations, comments like the following were received:

> "1978 projects will have a golden opportunity to do longitudinal studies of children and teachers that can be tied to demographic data and that your computers now provide a technology that can translate programmatic evaluation into a vehicle which can really assist schools in finding out what they are doing.

The complexity of the problem of managing and evaluating was described succinctly by one director:

> It is very frustrating because project personnel aren't really committed to evaluation or management efforts and view them as 'layons.' There are not many university faculty members interested or competent in project evaluation of the kind we need. Managing

> the political or collaborative aspects of program development drains the energy of persons trying to develop and deliver instruction . . . In a highly interactive system it is hard to document decision making processes.

The need for an early start in planning and carrying out the evaluation was cited again and again. The problems of design, of tailoring, of formative evaluation of the realities of the projects, of obtaining trained personnel who can manage complex data comfortably and of translating the results for a multiple clientele emerged again and again as persistent problems which need to be approached at the beginning and which in many cases will require technical assistance.

Some of the directors felt that Teacher Corps guidelines could help bring people to the point where they would be willing to accept evaluation if only through requirements. It was suggested that Teacher Corps should: "insist the guidelines attempt to measure gains for students in site schools"; "insist that evaluation efforts be supported by the grant including the best possible assistance"; and "include training for project and site personnel on the acceptance of evaluation efforts and the best possible design strategies."

Again, these comments are from a project which has a relatively complex design and whose director is obviously feeling the frustration of working almost alone to handle the design and political aspects simultaneously.

Time and money were mentioned frequently. Obviously many projects were trying a more complex evaluation design than could be supported by the resources they had.

> "The major problems are centered around the lack of resources to complete the job of evaluation as defined in the original proposal."
>
> "I believe we have finally worked things around so at least we have what we need to complete the job."
>
> "Our biggest lack was in the area of personnel."
>
> "It was only after a good deal of negotiation that we were able to secure two half-time research assistants. Somehow people underestimate the amount of time it takes to collect and process data."

The problems of identifying outside evaluators and of finding internal personnel who could do an adequate job were

mentioned frequently. Many feel that a full-time evaluator would be desirable, but the resources do not permit the necessary number of personnel to be allotted to the evaluation function.

Despite the difficulties of obtaining them, many of the projects supported the idea of outside evaluators.

Some of the projects which feel fairly comfortable with the evaluation process made comments revealing of the elements to be concentrated on.

> We have not really experienced a lot of difficulty with evaluation efforts. Some of the obstacles, however, have been coordinating the efforts into a consistent and logical approach, some resistance from participants (minor and overcome), the inevitable changes in project activities not originally envisioned which demanded the revision of objectives and the evaluation of outcomes, and getting reports out on time so as to have some effect and provide some input into further planning efforts. It should be added that these obstacles for the most part were overcome. They simply required more effort than originally planned for.

The call for clarity from all projects was consistent.

> The message is clear that evaluation, documentation and demonstration is needed, even required. Guidelines are needed in terms of the recommended expenditure, kind of evaluation desired, staffing for evaluation (who can do it, who should), and clear realistic reasons for the extensive documentation that is now being encouraged. Ones that will make sense to participants and staff alike. One has the feeling that Teacher Corps feels that evaluation and documentation are an intrinsic good; perhaps the rationale has to be reworked and/or restressed. We realize Teacher Corps is trying to fulfill each of these expectations already, but it is our impression that there is more effort needed in the above areas.

Several projects even suggested that part of the evaluation be organized around the same or similar instruments so that data can be comparable.

The involvement of the community and the assessment of the impact on the community also emerged from time to time as a troublesome theme. The necessity of tailoring evaluation to the social and political context was mentioned a number of times.

> Project evaluation *must* take into account the milieu in which the project functions . . . since the stage of the development of the society or culture with which the project is working, and how the perspectives of the members of the society bear upon all project operations from governance to curriculum development to political activity must be considered by the staff and institutional interests, must also be considered by evaluators and evaluation designs. Public involvement, and public support, and whatever terms the project's public sees fit to articulate is key measurement, perhaps the most important. Look to the clientele, or those served, rather than the servers even though opinion or vague senses are hard to quantify.

Results From Analysis of the Documents.

In general, the projects attempted to comply with our requests for material from their reports, contracts with evaluators, planning documents, and other material that would give us a picture of their evaluation structures.

These documents were read and a typology of the approaches to evaluation was gradually developed. In the case of 84 of the projects, the documents were complete enough that a classification could be attempted.

A two-way classification was devised according to the orientation of the evaluations (toward process only, toward product only and toward process and product) and the structuredness of the approach to evaluation (completely unstructured, structured/simple, structured/complex). Table 3 presents the results of that classification.

The projects which focused on process only tended to be somewhat more unstructured in their evaluation procedures, although fourteen of them used what we categorized as structured and complex approaches. In contrast, forty of the forty-seven projects which focused on process and product tended to use relatively structured and complex designs.

There is little evidence from the documents that the fourteen projects which focused on process only and used completely unstructured means of evaluation have, in fact, put together a functioning design.

Fifty-four, or nearly two-thirds, of the projects whose documents were analyzed used relatively structured and complex approaches to the evaluation process. This is surprising, given

TABLE 3

Classification of Evaluation Components of Eighty-Four Projects by Focus and Complexity

(Frequency and Per Cent of Total)

	Focus			
Complexity	Process	Product	Process and Product	Row Total
Completely Unstructured	14 (16.6%)	0 (0.0%)	0 (0.0%)	14 (16.6%)
Structured-Simple	9 (10.7%)	0 (0.0%)	7 (8.3%)	16 (19.0%)
Structured-Complex	14 (16.6%)	0 (0.0%)	40 (47.6%)	54 (64.2%)
Column Total	37 (44.0%)	0 (0.0%)	47 (55.9%)	84 (100%)

It can be seen that none of the projects attempted product-only evaluation, that is none went directly to the assessment of pupil learning or to an assessment of the effects on teachers without tracking the events that might lead up to an impact on other teachers or students.

We regard the absence of frequencies in the cells entitled "product only" as desirable in that product-only assessment does not permit inferences about the effects of project activities.

Several of the projects reported, and planning documents substantiated, that there was pressure from some school district personnel to invest the evaluation resources in product-only analyses. This pressure was due primarily to the current interest in the "bottom line" and the lack of interest in the dynamics of educational change.

Forty-four per cent of the projects concentrated largely on process and fifty-six per cent on process and product.

the relatively limited amount of resources available for evaluation and the considerable energy required.

The evaluation plans were also examined for evidence of the extent of their development in four categories:

The plan is clearly partial and incomplete.
The plan is relatively complete but there is little evidence of development.
The plan is relatively complete and there is evidently partial development with two categories: data use planned and data presently being used.
The plan is complete with evidence of substantial development and implementation.

Table 4 presents the results of this classification for the projects which maintained a process orientation and those representing a process/product orientation.

As might be expected, those projects with the process orientation which used the more formal approaches evidenced less development and implementation. Partly this is because the informal approaches themselves do not result in the same degree of tangible evidence about implementation. If there are informal interviews which are then discussed by a staff or other planning group, transmission of information is more likely to be oral and less likely to be written.

About a third of the process-only oriented evaluation plans gave evidence of relatively complete planning with substantial development and implementation.

Nineteen of the thirty-seven process-only projects (slightly over half) evidently have developed relatively complete plans and are presently using data.

Of the forty-seven projects with a process/product orientation and using a formal/complex approach, eighteen (almost 40 per cent) manifested relatively full development of plans with strong evidence of implementation of at least major components of the evaluation process. These represent the projects which have taken the most complex approaches and manifest the greatest degrees of program actualization. In addition, eighteen more of those projects have a relatively complete plan with partial development and nine of them are manifesting

TABLE 4

Classification of Evaluation Components of Eighty-Four Projects by Extent of Development and Implementation According to Focus and Complexity of Degree of Development
(In Frequency)

	Process Orientation				Process/ Product Orientation			
	Unstructured Approach	Structured-Simple Approach	Structured-Complex Approach	Row Total	Structured Simple	Structured Complex	Row Total	Grand Row Total
Plan Partial and Incomplete	10			10	3	4	7	17
Plan Relatively Complete: Little Evidence of Development	2	2	2	6		4	4	10
Plan Relatively Complete: Evidently Partial Development								
Data Use Planned	2			2	2	7	9	11
Data Being Used		5	2	7	2		9	16
Plan Complete: Evidence of Substantial Development and Implementation		2	10	12		18	18	30
Column Total	14	9	14	37	7	40	47	94

clear evidence of using data at present. *Thus, those projects which chose the more complex alternatives apparently made the greatest investment not only in planning but also in implementation.*

Although there is a continuum of levels of development evidenced in the documents, it would appear that the stronger and more comprehensive planning efforts are accompanied by the greater degrees of developmental effort later.

The data were analyzed to compare projects which did not follow an earlier project in the school district and higher education institution with those which did. Specifically, Cycle 11 projects which followed Cycle 9 projects were compared with those which had not been so preceded. In addition, Cycle 12 projects which followed Cycle 10 projects were compared with those sites where there had not been an immediately preceding cycle.

The differences were pronounced, especially for the Cycle 12 projects which were in their first year in the course of the present study. Projects with experience in the immediately preceding cycle were twice as likely to have generated complex evaluation designs and to have implemented them to the point where there was present use of data.

The strength of the association between prior experience and the complexity of the planning and developmental work in evaluation is striking. Because of the nature of the design of the present study we are unable to rule out a variety of hypotheses which compete with the proposition that experience and the ability to capitalize on previous ventures had a powerful effect on the evaluation in these projects. It is possible that it was the stronger sites which were able to compete successfully for a succeeding cycle, so that projects following previous cycles are generally stronger than others. It is also possible that the project staff persons associated with Teacher Corps become socialized to the points of view about evaluation which are disseminated by advocates of "management-by-objectives." Whatever the intent in Washington, this orientation is probably the field's image of what Washington regards as the best approach to management *and* evaluation!

The comments from project directors about the difficulties of formulating and implementing a successful evaluation pro-

gram would seem to indicate that time and experience are probably extremely important variables. The comments by new project directors, that the complexity of projects in the early years appears to be overwhelming, indicates not only the importance of project maturity, when many of the difficulties of "start-up" are overcome, but also the potentially important role that early technical assistance can play in helping people to solve the inevitable problems attendant to the project's beginning needs to be sorted out — how to aim and implement effective evaluation and, no doubt, training.

Table 5 indicates the kinds of data collection methods most frequently used in various dimensions or phases of projects.

The dimensions of the projects are here described as planning, implementation of training, the impact of training, the impact of teaching, and institutionalization.

Records of activities were used to document all dimensions (as were interviews and questionnaires). Altogether, descriptive records were by far the most frequent methods employed throughout most of the projects. A few projects used ethnographic records. When they did so, they tended to be across all the dimensions, although a few used ethnographic analyses of planning activities alone.

Behavior records of teachers were used in very few projects, but where they were used it was, appropriately enough, in the analysis of impact of training.

Ratings of teachers were used somewhat more frequently and also to assess the impact of training. Some use of pupil achievement data is apparently being made in about 25 per cent of the projects although only in about four or five projects is this regular and continuous. Training products and products generated by teachers in projects are evaluated to assess the implementation of training and also institutionalization.

This table, itself, indicates the potential complexity in designs. There are half a dozen cases where all of the cells marked with an X have been planned by a given project and, in fact, some data collection has been going on. There are a number of others where all of the cells except for those that intersect with ethnographic records are filled by some kind of developmental activity. Records, interviews, and questionnaires are used per-

TABLE 5

Data Collection Methods
Dimensions By Phase of Project

Dimensions	METHODS							
	Records of Activities	Interviews	Question-naires	Behavioral Records of Teacher Behavior (Direct Observation)	Ratings of Teachers	Pupil Achievement Tests	Products	Ethno-graphic Records
Planning	X	X	X					X
Implementation of Training	X	X	X				X	X
Impact of Training	X	X	X	X	X			X
Impact of Teaching	X	X	X			X		X
Institutional-ization	X	X	X				X	X

vasively throughout the dimensions of the programs in quite a number of cases.

This illustrates the point made above, that the process of evaluation can become pervasive. In one sense it is desirable to have embedded data collection, that is, to use records, interviews, and questionnaires, and other devices in the natural flow of activities in the project. But it is very easy to reach the point where such a practice makes evaluation extremely obstrusive. One very strong project is implementing an individualized instructional program of considerable dimensions. In that project pupil achievement data is collected continuously and fed back into the planning and implementation processes. The same project keeps complete records of the activities of staff, community members, teachers, and students, and conducts interviews and uses questionnaires to determine opinions of the success of activities.

The result is a very thorough-going process, one which would have been impossible had not the project (built on a very well-developed, modular, individualized curriculum for children) made the implementation of that curriculum its goal.

It is worthwhile noting that some projects used ethnographic records alone or combined with interviews to provide the staff an analysis of planning, implementation, impact and institutionalization.

With respect to most of the cells in this table, there are some projects whose personnel have considerable experience and could no doubt provide extensive technical assistance to others, were that desired.

Evaluation Issues in Teacher Corps: How Do Things Stand?

We now turn to the findings as they relate to the general issues, raised earlier, about evaluation. Teacher Corps data reported on evaluation are quite pertinent to each of the issues. Although the evaluation practices can be considered in terms of Teacher Corps purposes and configurations, Teacher Corps is a part of the larger educational enterprise, as illustrated by these results.

Breadth/Narrowness of the Concepts of Evaluation Being Used

As indicated earlier, the field of evaluation grew from an attempt to move away from the narrowly focused concerns of measurement *per se* into the broader questions involved in placing value on activities and alternative curricular patterns.

A large proportion of the Teacher Corps projects clearly attempt a broad range of documentation and assessment activities. In addition, devices are used to document program events, assess the opinions of personnel towards those events, to distinguish between planning and implementation, and attempt to monitor the effectiveness of both planning and training activities. Clearly Teacher Corps local projects are for the most part in the broader tradition of evaluation. This may not be good in many cases. It might be better to sharpen the designs, to make them highly focused and selective so as to reflect indigenous needs.

Formative and Summative Evaluation Compete for Emphasis

Overall, the emphasis throughout Teacher Corps is more formative then summative. Projects seem to be grappling with the problems of collecting data early enough in the program to track the emerging shape of events and use data to improve activities.

Perhaps here is where a division of labor might be achieved. It might be possible and advisable for local projects to devote themselves mainly to formative process evaluation, provided summative and product evaluation are conducted on a regional and national basis — the summative evaluation efforts could be coordinated across projects. Presently there is a great commonality of aims and effort among projects. Costs and results could be shared, thus reducing the burden on local projects and, with some common instrumentation, the power of the results could be greatly enhanced. Over five years, this would be much more possible than has been the case in the past.

Experience seems to indicate that a substantial amount of evaluation on a project should be done inside by those who have regular project responsibilities. Everyone, rather than someone, might be designated as evaluators. There is danger

lest what is considered to be everybody's business may become nobody's but it also is true that without widespread participation we suffer in many ways. There seems to be wisdom in having individuals or teams designated operational responsibilities, under guidance of a representative advisory group. Thus, the need inside is for both "experts" and generalists.

When well selected and rightly used, outside experts can have a useful impact, and they certainly contribute to the credibility of evaluation to policy makers, the public, and other interested groups. The question is how to use them well, for all too often they are not. We have some intuitive hunches, some are derived from experience, and the data hint at more propositions. But the matter needs much more thorough study than it has received, for the use of outside consultants has become widespread. We recommend this as a high priority for planning in 1978.

Process and Product Compete in Both the Formative and Summative Domains

In Teacher Corps the weight is on a combination of process and product. Where there is one, alone, it is process.

Given the participatory nature of the Teacher Corps programs, this practice appears to be entirely appropriate, for the broad base of participatory governance makes it difficult to separate process from product. It is no doubt true, also, that their intertwined nature is one reason why projects use a great deal of internal evaluation, for it would be very difficult for an outsider to learn enough about the totality of a project to provide a comprehensive view of events and their effects. In the future it might be advisable to increase the total amount of attention given to product evaluation to bring it into better balance.

The Virtues of Objective and Subjective Sources Are Debated Along with Goal-Oriented and Goal-Free Approaches

The objective-subjective debate does not seem to have much currency in Teacher Corps which appears to use (in nearly all cases) combinations of objective and subjective sources. While the theoreticians may argue the merits of one or the other,

these practitioners have generally recognized the advantages of both.

Generally speaking, approaches such as those advocated by Stufflebeam and Provus predominate, with the accompanying emphasis on all phases of project operation in attempts to identify linkages between planning and action components. Objective and subjective sources tend to be used through all phases. Nearly all evaluation plans are goal-oriented. Goal-free and "connoisseurship" approaches, suggested by Stake, Scriven, and Eisner, receive minor attention, although they are beginning to be discussed.

Based on our analysis of the 11th and 12th Cycle activities, Teacher Corps seems to follow the modal or most popular current ideas in the field of evaluation.

The Relative Merits of Periodic, Continuous, and Post-Project Plans are Argued

None of the projects have adopted post-project plans alone, although about 25 per of the projects are attempting to develop a summative picture in terms of changes in teachers or pupils. The flow is definitely toward relatively continuous data collection and analysis, although some projects have opted for periodic assessment. The periods, generally, however, are closely tuned to the times when phases of the projects are initiated or concluded, and thus even the periodic assessment plans have the flavor of continuous ones.

As indicated throughout the report, the attempt to build formative, process and product-oriented, comprehensive evaluation programs using objective and subjective sources with data collection on a continuous basis is extremely difficult. This is clearly the primary explanation of the types of frustrations and concerns voiced by project directors. The approach which Teachers Corps has taken to evaluation requires a high degree of cooperation among all project personnel so that each person sees evaluation as a part of the base of project operations. It retains a flavor of an "add on," "something being done to us" operation. It is extremely difficult to govern and conduct evaluation so that participants share in its ownership.

This is an important problem to approach. While no easy

solution exists, help could come from those projects which have at least partially solved the problem. It is likely that models can be generated which can help projects approach the question more effectively in the future. *In addition, the resources which have been allotted to evaluation are in most cases so slender as to preclude the level of support which the evaluation aspires to.* This probably is the clearest lesson that emerges from this report. The alternatives are clear: lessen the degree of the effort, increase the resources, or find a more efficient way to evaluate.

On-Site and Off-Site Methods Are Compared for Cost and Fidelity

This is really not much of an issue at the local level where evaluation is either designed by the staff or commissioned by them. No fully off-site evaluations were detected.

In a few cases, however, outside evaluators, even though they may have visited the project, have behaved as if they were quite a distance away, using questionnaires which of necessity have had to be designed rapidly. Although it is not possible, of course, to tell precisely, a number of the questionnaires which were included with the materials did not "Feel" attuned to project realities as we know them. Some external evaluation simply bombed out.

It is possible for an inside or outside evaluator to act as if he/she were at a great distance. Clearly the issue is live at the local level. Even the staff of the project reports that in many cases there is difficulty in getting school personnel to accept the idea of evaluation itself and also the legitimacy of the instrument which has been devised, even by persons quite close to the project.

Sponsorship and Ownership Entail Substantive as Well as Political Questions

Although many projects have responded more than cooperatively to Teacher Corps/Washington's request for evaluations, it is clearly difficult for the "spirit of evaluation" to pervade all project personnel from all of the cooperating institutions. It is not the most popular of Teacher Corps activities. Evaluation in many cases still has a slightly unnatural feel.

What do these findings mean for the future? Where are the problems and difficulties, the needs, the places for potential new directions, and how can a better distribution and division of labor be achieved?

Many large scale nation-wide programs in education have been widely criticized because evaluation efforts at project sites have been weak and ineffectual. Why should not this be the case in Teachers Corps? The picture that emerges from these data, however, is one of surprising strength in many cases. While there is considerable variation in evaluation efforts amongst projects the central tendency is that of a substantial, serious, effort in evaluation that gives no cause for shame.

The evidence is that many of the Teacher Corps projects have devoted considerable energy to developing quite complex evaluation plans and are grappling with contemporary evaluation issues at a level equal to the better thinking in evaluation circles today. Many of these imaginative and effective evaluation activities ought to be shared widely both within and outside of Teacher Corps. This is not to say that everything is as good as it should or could be. Hence, we herewith summarize some of our judgments in a format organized around *problems* and *recommendations* to consider for the future.

Problem #1

A major recurring problem is an attempt to use elaborate theoretical models that outstrip available resources and may be beyond the scope or need of a local project.

There seems to prevail in altogether too many places the belief that an elaborate management by objectives, performance based and discrepant analysis model is necessary which requires and results in an inordinate amount of paper with very little on it. Current concern for conservation of our natural resources alone should bring this practice to a halt.

The attempt of projects to touch as many bases as possible in the evaluation process and in many cases to attempt to implement the management by objectives philosophy results in ex-

tremely long lists of events to be studied. Especially when they try to follow Provus's or Stufflebeam's models the lists can become very long indeed in an attempt to document aspects of the planning process, training activities planned and implemented, the responses of personnel to those activties, etc.

Given the current limited investment in evaluation, the projects need to rely on self-report data for the most part, resulting in the extensive use by participants of logs, diaries, "activity check lists," and questionnaires about activities. While at best these can be useful in presenting a comprehensive picture of the projects, they can, and in many cases do become loose collections of relatively superficial data, of little use and especially irritating to those who are required to generate them.

The evaluation ideas of Stufflebeam and Provus are the most widely prevalent of any of the currently prominent theoretical conceptions in use in 11th and 12th Cycle Teacher Corps projects. These conceptions are quite ambitious. Generally speaking, they aim to document almost everything that happens in the course of the project, at least to sample randomly or selectively from major variables within a project. It is easy to follow the formulae of these conceptualizations and develop a pattern which is more complex than the project itself. Although the nature of action projects is complex in a fundamental way, the projects themselves are relatively small. But we noted in some of the documents submitted an apparent, almost compulsive, attempt to follow out the schemes through to their last and most minute detail, which seemed unnecessarily burdensome as well as useless.

Why are these systems so widely used? Are they simply the best available in the profession, having won out in the honest and open competition in the market place? Have they been "marketed" or "sold" by some central and persuasive salesman? Have they inadvertently or unintentionally received the blessing of powerful authority figures? Whatever the reasons, problems in their use appear to have occurred. These may be due to the schemes themselves, for which possible remedies may be available. Or more optimal approaches may have been developed which may more closely meet the needs of a program such as Teachers Corps.

Recommendation #1

Our first recommendation is to make sure that projects are not rewarded for developing elaborate schemes of evaluation, and that more simplified evaluation plans be encouraged.

The word from Washington needs to be loud and clear that there is no anointed model (certainly not the ones currently in most common use), and that the power rather than the length of the scheme is most important. Variety, however, is needed both in the technical assistance and in the the documents provided Teacher Corps. Evaluation models should be distributed by Teacher Crops/Washington that describe available options and each option should have the Surgeon General's label—*Danger, open and use carefully—This may be dangerous to your health.*

It is advisable not to collect structured data where unstructured data will suffice. For example, if one has planned only six activities, it is not necessary to use questionnaires to determine whether they were implemented. A staff meeting can provide adequate data on whether something did or did not happen in the case of many kinds of activities. We found a number of complex "documentation only" efforts which used relatively loose opinionnaires with simple statements about whether activities did or did not occur. A good manager can obtain much of this information simply through his own eyes and ears, making it unnecessary to formalize so many of the aspects of documentation.

A few projects focused carefully on selected aspects of their programs and attempted to collect reliable data in depth on those limited portions of their programs. This is preferable to obtaining large quantities of superficial data.

Problem #2

In many cases the intitial proposals contain over-elaborate designs that later must be cut to realistic proportions.

This springs partly from trying to "read" what is wanted in Washington and partly from a seemingly inexorable force inherent in proposal writing, namely to err on the side of inclusion, sweep every potentially good idea in and weed out—later.

Recommendation #2

We recommend that Washington be quite clear about what it does not want as well as what it does in the valuation prose of the proposal.

A constructive step was taken in the guidelines for 1978, limiting the proposal to not more than 100 pages, which increased the probability that there would be something of substance on each one.

More specific examples of do's and don'ts could be included in the guidelines. For example: "State your goals and objectives clearly but succinctly. Give an example of one or two as illustrative of how the others are to be done."

As a second part of this suggestion, Washington could, in the negotiating phase, require that plans be tailored to realistic proportions, and build in guarantees as much as possible that plans will be followed.

Problem #3

The evaluative conceptualizations are often too loosely conceived, attempt to cover too many bases, and do not enable the project to analyze its overall strategy.

Recommendation #3

We recommend that the guidelines for the project, at the proposal writing stage, be written so as to encourage tighter conceptualization of evaluation efforts.

This would enable projects to better examine significant variations. For example, degree of participation can be usefully correlated with degree of involvement of training. Degree of integrativeness in planning and carrying out activities can be looked at in relation to satisfaction and the transfer of the effects of training to different arenas in the school.

It is possible for evaluation designs to be made both more elegant and more parsimonious if they state explicitly some of the assumptions—perhaps even the theories—which guide programs and then test those assumptions. Simply finding out what has happened and what effects it has had is useful. To test the assumptions of a project is even more useful. To build the

evaluation so that it provides a reflector to the project about what is happening, combined with the testing of assumptions, would be the most useful of all.

However, to provide for a reflective function probably means contracting outside for at least that aspect of evaluation. A project that plans to reflect on itself will not be able to go much beyond its original capability for reflection. The outsider can be useful at that point. The broad perspective, the relative impartiality and objectivity of the "outside" view and the greater experience that may be represented, can be helpful in enabling the project personnel to carve out time from pressing schedules and deadlines, and to be reflective in response to the queries raised by the outside evaluators. Under current conditions this highly desirable state of affairs is too seldom realized and may be one of the reasons for the unduly negative attitudes that tend to surround evaluation among many participants.

Problem #4

The fundamental task of assessing pupil achievement and relating it to teacher behavior remains in a relatively undeveloped and primitive form, much more so than is necessary given the present state of the art.

Only 30 per cent of the projects that returned the questionnaires indicated that they were using pupil achievement "quite a bit" or "extensively" as a focus in their evaluation, and 70 per cent of the responses indicated little or no such use. A somewhat larger number (70 per cent) were attempting to document teacher and intern growth, but mostly by the means of ratings, with few using the "hard," more reliable data derived from observation. Surprisingly, given the relatively greater availability of valid and reliable achievement instruments, as contrasted with attitude measures, attitudes are more attended to (67 per cent "quite a bit" and "extensively") than learning (30 per cent).

The assessment of these outcomes is most difficult but there is much that is known that needs to be applied.

Recommendation #4

The recommendation to overcome these shortcomings is to devote more attention to the problems of measuring pupil and

teacher growth, plan earlier and more realisticially for assessing these outcomes, and provide more technical assistance.

If summative evaluation plans that include changes in teacher behavior and associated changes in pupil behavior are considered to be desirable, it should be made clear from the beginning that this is policy and that considerable technical assistance will be available. Relatively few projects use the more precise and powerful objective methods now available for assessing teacher behavior.

Without assessing such changes it is difficult to make inferences about project effects on children. Yet, given the number of Teacher Corps personnel already engaged in evaluation activities, it would be possible to offer training in design and techniques to persons in local projects who can accept such training and use it fruitfully. Local persons in the National Urban/Rural Program easily learned the more objective techniques and used them effectively. Generally speaking, it should be easier to prepare Teacher Corps personnel for the more technical tasks than it was in the case of Urban/Rural, because Teacher Corps includes a much higher proportion of professionally trained personnel.

The attempt to build effective summative evaluation plans requires that data be quickly gathered and fed back. Such a system takes considerable effort to develop in an ongoing project. Yet, the techniques are relatively standard. If projects are to be encouraged to engage in relatively formal evaluation schemes, such models should be provided relatively early. There should be sufficient development at the national level so that a transportable technology can be provided to local projects.

An alternative to be considered is whether local projects should be required or even encouraged to do summative evaluation on their own. (See the National Task Force suggestions that there be a division of labor, with national resources and efforts devoted to the summative evaluation tasks).

A great many efforts to collect data in the pupil achievement domains are necessarily summative, given the relatively short time of the project. Exceptions are those projects which are using one of the highly precise systems-organized types of in-

struction with children where each unit or module of student activity is accompanied by built in evaluation. It is encouraging that no projects attempted to look at pupil achievement in a summative way without also looking at process. It would be useful in the future to approach the technology of utilizing achievement data in a formative way as well. For example, some insightful designs in research on teaching suggest that a new class of variables—pupil behavior—might well be looked at, in order to make possible the linking together more closely of teacher behavior and pupil achievement, as shown in Figure 5.

Probably one of the most important decisions is going to be in the area of costs and benefits. The major decision, as we see it, involves whether to encourage an "evaluation spirit" throughout project staffs so that engaging in some evaluation activities continuously as a part of one's work and as a client of training is looked upon as an ordinary and acceptable activity. If this spirit can be accepted, then providing greater local participation in evaluation design and hence increasing the ownership of evaluation can be accomplished. In addition, information from participants can be gathered more easily and all project personnel can become accustomed to a regular feedback process and to the utilization of data for continuous program improvement.

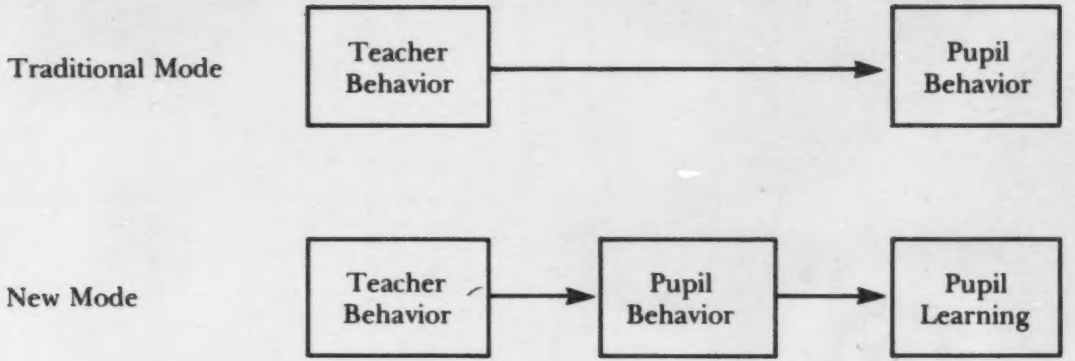
We believe that this is the optimal course of action. Otherwise, evaluation should probably be extremely selective and quite narrowly targeted.

Problem #5

Our pervasive impression is that evaluation is not a very well liked part of the program in many sites. There is a considerable amount of grumbling about it. This is not unique to Teacher Corps, however. It is endemic in education, and indeed throughout society when evaluation activities are undertaken.

This negative case is unlikely to be fully eradicated, but the necessity for evaluation is not likely to go away either. Hence, it is incumbent upon us to see that evaluation be done in as reliable, valid, credible, and efficient a manner as possible, and to eliminate as much of the needless noise as possible.

Figure 5. Illustration of New Way of Looking



Recommendation #5

Can anything be done to make the process more palatable? The most likely strategy, we believe, would be to make evaluation so much an integral part of the project that it becomes almost indistinguishable. Stop having it be something tacked on.

This would also mean reducing at the local project level the amount of "external" and unrealistic summative evaluation demands that often cannot be well executed anyway. It would help to provide competent resources, both dollars and technical help, in a liberal fashion, perhaps on a matching basis, to ensure what was consumed was what was perceived to be needed. There would need to be some variety in the types of technical assistance provided so that projects would not feel that they were being pushed into a particular mold. Making evaluation more palatable would also be achieved by overcoming as many of the shortcomings herein enumerated as possible.

Two methods predominate for documenting changes in teacher behavior. One of these is questionnaires to the teachers themselves asking for self reports about the effects of training. The other is rating scales, usually global in nature, where instructional personnel or Teacher Corps staff or public school administrators and supervisors observe and rate teaching performance. Relatively few projects use the more powerful method of systematic observation with trained observers who have achieved reliability in the recording of behavior measurement of the manifestation of demonstrated teaching skill.

Self-reports are useful because they actively engage the trainee her/himself and, generally speaking, self reports correlate fairly well with certain kinds of behavioral descriptions. However, global ratings, whether by self or others, i.e., general ratings of teacher effectiveness, are not particularly reliable and especially tend to differ across role groups.

Since the delivery of inservice activities is the focus of most of the Teacher Corps investment, it is pivotal that an effective evaluation provide a careful documentation of the effects on teacher behavior. Without such a careful documentation it is extremely difficult to draw inferences about the effects of the

planning process and social context of the project and virtually impossible to make any inferences about the causation of changes in pupil behavior.

Problem #6

There appears to be some unnecessary duplication of effort, e.g., in construction of instruments, in conducting summative program evaluations.

Recommendation #6

We recommend dividing up the labor and improving communication.

The aim here is not to destroy local initiative in instrument building, but to promote a greater sharing. Instrumentation frequently requires considerable time and skill. Networks and Teacher Corps/Washington could undertake the preparation, circulation, and keeping up-to-date, of an evaluation instrument source book—looseleaf—that could become a usable part of the shelf of every project. Not everything produced and sent in by every project should be included. There would need to be selection and careful organization of the manual so that project members could readily find what they need at a particular moment.

On the summative evaluation side, it might be desirable for Teacher Corps/Washington, in cooperation with projects in the field, to formulate a policy statement about summative evaluation efforts. This would enable assumption of greater responsibility in the region and nationally, and call for local participation on a more limited but specific basis.

The effort required to design and conduct an effective summative evaluation and the time span necessary to carry it out puts it beyond the reach of local projects in many instances. Greater resources at the national level could substantially upgrade this type of evaluation, which in turn would make the results more useful to policy makers at both national and local levels.

Problem #7

While there is a good "family feeling" in Teacher Corps, and

there is considerable sharing at regional and national meetings, there are still too many untapped resources in the projects that need to be drawn upon in providing technical assistance in many needed areas.

Recommendation #7

Our final recommendation, which we make strongly, is for the expansion in all areas, including evaluation, of technical assistance arrangements both in dollars and in people. The latter should be carefully and more liberally selected from the projects.

Examples of needs for which expertness already exists include: how to involve local people; how to get school and community to work together with genuine parity; how to obtain genuine community participation, and how to use electronic data storage and retrieval and other data processing systems economically and efficiently.

It appears that there are a number of steps which can be taken as the 1978 cycle begins and a number of important decisions which should be made regarding the messages on evaluation which will come from the Washington office.

A Concluding Note

This survey of evaluation practices in the current two-year cycle projects (11th and 12th) of Teacher Corps was undertaken mainly to learn better how to proceed with evaluation in the new five-year projects beginning in 1978. It is to this end that the interpretations and recommendations have been addressed. We hoped that some of the information would be helpful during the year remaining in 12th cycle projects. The question arises of how much of what we have learned may be equally applicable to both the five and the two year programs? For the most part, we believe that the interpretations and recommendations are applicable to both. There should be somewhat greater opportunity to overcome some of the shortcomings that have faced evaluation in the past with the longer time, more money, and larger scope of the program beginning in 1978. For example, with a year to plan, evaluation schemes can be more sharply focused and better designed. Better longitudi-

nal studies of the effects of the program both on students and educational personnel will be more possible within a five year than within a two year period.

The time for a powerful and persuasive national design for evaluation in Teacher Corps has come. It should be closely coordinated with local efforts in such a way as not to be burdensome and threatening to local projects but, nonetheless, be designed to satisfy the many different audiences of policy makers and the public as well as those within the profession who need and consequently will use the results. This is a big order but not an impossible one. The amount of effort required of Teacher Corps/Washington in evaluation matters should probably be greater than it has been in the past, but so will be rewards. This is neither the time nor the place to spell out the details of a national design, but many of the ideas are contained in the Task Force Report on Evaluation and in this report. It is anticipated that much required to mount new effective evaluation efforts may be incorporated in the forthcoming request for a proposal in the field of evaluation by National Teacher Corps and that that effort will propel us along the way.